

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 10:27:46 ; Search time 1299 Seconds
(without alignments)

11496.806 Million cell updates/sec

Title: US-10-028-248A-210

Perfect score: 3260

Sequence: 1 ggcacgagcaggagcttccc.....aaaaaaaaaaaaaaaaaaaa 3260

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3259	100.0	3260	16	US-10-028-248A-210
2	3259	100.0	3260	16	Sequence 210, App
3	3232.6	99.2	4576	9	US-10-107-782-210
4	3232.6	99.2	4642	15	Sequence 24, Appl
5	2999.8	92.0	8444	16	US-10-133-172-3
6	2999.8	92.0	8444	16	Sequence 1, Appli
7	2995.6	91.9	8495	16	US-10-028-248A-3
8	2995.6	91.9	8495	16	Sequence 3, Appli
9	2853.2	87.5	4962	15	US-10-133-172-19
10	1936.8	59.4	4706	15	US-09-842-930A-1
11	1936.8	59.4	4706	15	Sequence 1, Appli
12	708	21.7	718	15	US-10-133-172-1
13	537	16.5	722	9	US-10-023-896-38
14	424.6	13.0	3681	16	Sequence 38, Appl
					Sequence 559, App
					Sequence 1320, Ap

Sequence 30654, A
Sequence 610, App
Sequence 1824, Ap
Sequence 12, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 16736, A
Sequence 230, App
Sequence 5045, Ap
Sequence 2410, Ap
Sequence 8459, Ap
Sequence 32777, A
Sequence 1252, Ap
Sequence 58386, A
Sequence 58386, A
Sequence 446, App
Sequence 124, App
Sequence 1, Appli
Sequence 291, App
Sequence 1020, Ap
Sequence 445, App
Sequence 15, Appl
Sequence 880, App
Sequence 18, Appl
Sequence 293, App
Sequence 18, Appl
Sequence 17767, A
Sequence 44, Appl
Sequence 1663, Ap
Sequence 13589, A
Sequence 488, App

408.8 12.5 448 10 US-09-918-995-30654
391.2 12.0 2863 16 US-10-104-047-610
285 8.7 2160 16 US-10-104-047-1824
17 259.6 8.0 2753 10 US-09-774-639-12
18 259.6 8.0 2753 10 US-09-969-730-13
19 259.6 8.0 2753 17 US-10-621-363-13
20 259.6 8.0 2753 10 US-09-918-995-16736
21 247.4 7.6 395 10 US-10-305-720-230
22 187.4 5.7 1482 16 US-09-783-590-5045
23 176.2 5.4 293 9 US-09-918-995-2410
24 143 4.4 473 10 US-09-918-995-8459
25 135 4.1 420 10 US-09-918-995-32777
26 125.2 3.8 474 10 US-09-918-995-1252
27 107.2 3.3 409 9 US-09-833-381-1252
28 82 2.5 466 13 US-10-085-783A-58386
29 82 2.5 466 16 US-10-242-535A-58386
30 81.4 2.5 1144 14 US-10-044-090-446
31 81.4 2.5 1144 15 US-10-247-671-124
32 81.4 2.5 1414 9 US-09-799-118-1
33 81.4 2.5 1422 16 US-10-295-027-291
34 81.4 2.5 1430 16 US-10-295-027-1020
35 81.4 2.5 1728 14 US-10-044-090-445
36 81.4 2.5 7840 12 US-10-634-574-15
37 78.2 2.4 351 15 US-10-060-036-880
38 78.2 2.4 1414 16 US-10-133-937-18
39 78.2 2.4 1414 16 US-10-295-027-293
40 78.2 2.4 1414 16 US-10-159-563-18
41 74 2.3 396 10 US-09-918-995-17767
42 73.6 2.3 1411 9 US-09-765-231A-44
43 60.4 1.9 2627 16 US-10-108-260A-1663
44 60 1.8 60 10 US-09-908-975-13589
45 58.8 1.8 781 13 US-10-276-774-488

ALIGNMENTS

RESULT 1

US-10-028-248A-210
Sequence 210, Application US/10028248A
Publication No. US20030235882A1
GENERAL INFORMATION:
APPLICANT: Shimkets, Richard
APPLICANT: Patturajan, Meera
APPLICANT: Vernet, Corine
APPLICANT: Casman, Stacie
APPLICANT: Malyankar, Uriel
APPLICANT: Shenoy, Suresh
APPLICANT: Spytek, Kimberly
APPLICANT: Gangolli, Esha
APPLICANT: Miller, Charles
APPLICANT: Boldog, Ferenc
APPLICANT: Li, Li
APPLICANT: Li, Li
APPLICANT: Taupier Jr, Raymond J
APPLICANT: Kekuda, Ramesh
APPLICANT: Smithson, Glennda
APPLICANT: Zerhusen, Bryan
APPLICANT: Liu, Xiaohong
APPLICANT: Colman, Steven
APPLICANT: Tchernev, Velizar
APPLICANT: Si, Jingsheng
APPLICANT: Edinger, Shlomit
APPLICANT: Stone, David
APPLICANT: Sciore, Paul
APPLICANT: Millet, Isabelle
APPLICANT: Rothenberg, Mark
TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
TITLE OF INVENTION: Thereof
FILE REFERENCE: 21402-222
CURRENT APPLICATION NUMBER: US/10/028,248A
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/256619
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 60/262959

;
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 210
; LENGTH: 3260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3171)
; OTHER INFORMATION: Wherein n is A, C, G, or T
; US-10-028-248A-210

Query Match 100.0%; Score 3259; DB 16; Length 3260;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGACGAGGAGCTTCCCAAGAACCCGAAACCTTCCAGTATTTCTTCCAGTTGCAGG 60
Db 1 GGCACGACGAGGAGCTTCCCAAGAACCCGAAACCTTCCAGTATTTCTTCCAGTTGCAGG 60

QY 61 AGCAATTCGTAAGAATCTGGTCCGCGCCAGGCCCTTCACTGTTTTCACCTTTATCTG 120
Db 61 AGCAATTCGTAAGAATCTGGTCCGCGCCAGGCCCTTCACTGTTTTCACCTTTATCTG 120

QY 121 CAGCCTTTTCATAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGG 180
Db 121 CAGCCTTTTCATAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGG 180

QY 181 TTCTTCGGTACCATTGTTGCTCGCTCCACCAGCTGCTTCTGGAACCTTGAATTTGATCT 240
Db 181 TTCTTCGGTACCATTGTTGCTCGCTCCACCAGCTGCTTCTGGAACCTTGAATTTGATCT 240

QY 241 CAAATGCTACTTCCCTCCAGGAGACCAATAGTCAATCTCGGCTCTCCAGAGCACGGTGT 300
Db 241 CAAATGCTACTTCCCTCCAGGAGACCAATAGTCAATCTCGGCTCTCCAGAGCACGGTGT 300

QY 301 ATATAAATAAAGGCTAAGATCATATCCAGTGATATCATCCAGTGATATCATCTAATGGATTGTTTC 360
Db 301 ATATAAATAAAGGCTAAGATCATATCCAGTGATATCATCCAGTGATATCATCTAATGGATTGTTTC 360

QY 361 ATATCATGACAAAATTGCTATCTCCAAAAATTGCTTATCATCTCCCAAGACAACTCTG 420
Db 361 ATATCATGACAAAATTGCTATCTCCAAAAATTGCTTATCATCTCCCAAGACAACTCTG 420

QY 421 GAAGAATCTGCATAAATCTTAGCTTTGGCAACAAATGCTGATCAATCAATTTAGCA 480
Db 421 GAAGAATCTGCATAAATCTTAGCTTTGGCAACAAATGCTGATCAATCAATTTAGCA 480

QY 481 ACTTAATACAGGACTCAGGTTTGTGCTGAGTGTCTATCCAGATCCCATCCACCCAGTCA 540
Db 481 ACTTAATACAGGACTCAGGTTTGTGCTGAGTGTCTATCCAGATCCCATCCACCCAGTCA 540

QY 541 CTCTCTTCTGGCCCAACGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600
Db 541 CTCTCTTCTGGCCCAACGACCAAGCCCTCCATGCCCTACCTGCTGAACAACAGGACTTCC 600

QY 601 TGTTCACCAAGACCAACAGGACAAAGCTGAAGGATTTTGAAGTTTCAATGATACGAG 660
Db 601 TGTTCACCAAGACCAACAGGACAAAGCTGAAGGATTTTGAAGTTTCAATGATACGAG 660

QY 661 ATGCCAAGGTTTTAGCTGTGGATCTTCCACATCCACTGCCCTGGAAGACCCCTGCAAGTT 720
Db 661 ATGCCAAGGTTTTAGCTGTGGATCTTCCACATCCACTGCCCTGGAAGACCCCTGCAAGTT 720

QY 721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTGTAATGGCC 780
Db 721 CAGAGCTGAGTGTGAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTGTAATGGCC 780

QY 781 AAACCTGCAGAAATGTGCAGCGGAGCTCTTGTGTTGACCTGGGTGTGCGCTACCGCATTTG 840
Db 781 AAACCTGCAGAAATGTGCAGCGGAGCTCTTGTGTTGACCTGGGTGTGCGCTACCGCATTTG 840

QY 841 ACTGTCTGTGATGTGATCCCACTGGGGGCGCTGTGACACCTTTACTACTTTCGATG 900
Db 841 ACTGTCTGTGATGTGATCCCACTGGGGGCGCTGTGACACCTTTACTACTTTCGATG 900

QY 901 CCTCGGGGAGTGTGGGAGCTGTGTAATCTCCAGCTGCCAGTGGAGTGAATTAACCAA 960
Db 901 CCTCGGGGAGTGTGGGAGCTGTGTAATCTCCAGCTGCCAGTGGAGTGAATTAACCAA 960

QY 961 AGGGTGTGAAGACAGAAAGTGTCTCAAACTGCCCTTCAAGAGGAACCTCGAAGGCTGCC 1020
Db 961 AGGGTGTGAAGACAGAAAGTGTCTCAAACTGCCCTTCAAGAGGAACCTCGAAGGCTGCC 1020

QY 1021 GGGAGCGTGCAGCCTGTGTGATACAGATCCCAAGTGTGTCAGAGGCTACTTCGGGCGAG 1080
Db 1021 GGGAGCGTGCAGCCTGTGTGATACAGATCCCAAGTGTGTCAGAGGCTACTTCGGGCGAG 1080

QY 1081 ACTGTCCAGGCTGCCCTGGAGGACCAATGTCGCCGTGTAAATACCGGGGTGTCTGCCCTG 1140
Db 1081 ACTGTCCAGGCTGCCCTGGAGGACCAATGTCGCCGTGTAAATACCGGGGTGTCTGCCCTG 1140

QY 1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGACCGCGTGTG 1200
Db 1141 ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGACCGCGTGTG 1200

QY 1201 AGATGTCTGGCCGGGAGATTCGGGCTGANTGTCTGCCCTGTGCTGTGACAGCAACG 1260
Db 1201 AGATGTCTGGCCGGGAGATTCGGGCTGANTGTCTGCCCTGTGCTGTGACAGCAACG 1260

QY 1261 GACAGTGCATGATGTCATCAGGGCTCCGGCAGTGCCTCTGTGAAACGGGTTGACAG 1320
Db 1261 GACAGTGCATGATGTCATCAGGGCTCCGGCAGTGCCTCTGTGAAACGGGTTGACAG 1320

QY 1321 GCCCTCTGTGTGACACTCAGGAGTGTTCCTGTGACGTGTGACGCTCTTGTCTGTCTC 1380
Db 1321 GCCCTCTGTGTGACACTCAGGAGTGTTCCTGTGACGTGTGACGCTCTTGTCTGTCTC 1380

QY 1381 ATGCCACCTGTAAAGAGAAACAACAGTGTGAGTGTAACTGGAATTAAGAGTGAACGAA 1440
Db 1381 ATGCCACCTGTAAAGAGAAACAACAGTGTGAGTGTAACTGGAATTAAGAGTGAACGAA 1440

QY 1441 TCACATGACACAGTGTGGATTTCTGCAACACAGCAACCGGGGCTGTGCAAGGTTGCCA 1500
Db 1441 TCACATGACACAGTGTGGATTTCTGCAACACAGCAACCGGGGCTGTGCAAGGTTGCCA 1500

QY 1501 GATGTCCCAAGAGGCAAGAGTCTCTGACAGTGTCCAGAGGGAATCAAAAGGGGAGC 1560
Db 1501 GATGTCCCAAGAGGCAAGAGTCTCTGACAGTGTCCAGAGGGAATCAAAAGGGGAGC 1560

QY 1561 GGCACAGTGTGCACAGATAGACCCCTGTGACAGCGGCTTAAACGAGGTTGTACAGAGC 1620
Db 1561 GGCACAGTGTGCACAGATAGACCCCTGTGACAGCGGCTTAAACGAGGTTGTACAGAGC 1620

QY 1621 AGCCACCTGTAAAGATGACAGCGCGGCAAGCAAGTGTGAGTGAAGTCACTATG 1680
Db 1621 AGCCACCTGTAAAGATGACAGCGCGGCAAGCAAGTGTGAGTGAAGTCACTATG 1680

QY 1681 TCGGAGTGGCTGAATCTGTGAGCGGAGAGTGTCCCATTTGACAGGAC 1740
Db 1681 TCGGAGTGGCTGAATCTGTGAGCGGAGAGTGTCCCATTTGACAGGAC 1740

QY 1741 ATGGCAGTGTGCACAGCAACGCAATGTGTGACCTCCACTTCCAGATACCACTGTG 1800
Db 1741 ATGGCAGTGTGCACAGCAACGCAATGTGTGACCTCCACTTCCAGATACCACTGTG 1800

QY 1801 GGGTGTTCATCTAGCTTCCCTCCAGTGGGCCAGTATTAAGCTGACCTTTTACAAAGCCAGAG 1860

;; PRIOR FILING DATE: 2001-08-09
;; PRIOR APPLICATION NUMBER: 60/279,344
;; PRIOR FILING DATE: 2001-03-28
;; NUMBER OF SEQ ID NOS: 215
;; SOFTWARE: CuraseqList version 0.1
;; SEQ ID NO 210
;; LENGTH: 3260
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (171)..(2861)
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (3171)..(3171)
;; OTHER INFORMATION: Wherein n may be a, c, g or t
US-10-107-782-210

Query Match 100.0%; Score 3259; DB 16; Length 3260;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAACCTCCCAAGTATTTCTCCAGTTGACG 60
Db 1 GGCACGAGCAGGAGCTTCCCAAGAACCCGAAACCTCCCAAGTATTTCTCCAGTTGACG 60
QY 61 AGCATTTCTGGAAGATCTGGTCGGCCAGGCGCCCTTCACTGTTTGGCACCTTTATCTG 120
Db 61 AGCATTTCTGGAAGATCTGGTCGGCCAGGCGCCCTTCACTGTTTGGCACCTTTATCTG 120
QY 121 CAGCCTTTGATGAGGAGCTCGGGTTAAAGACTGGGCAAAATACGGTTTAAATCCCAAG 180
Db 121 CAGCCTTTGATGAGGAGCTCGGGTTAAAGACTGGGCAAAATACGGTTTAAATCCCAAG 180
QY 181 TTCTTCGGTACCATGTTGCTGCTGCCAGCTGCTTCTGGAACCTTGAATTTGATCT 240
Db 181 TTCTTCGGTACCATGTTGCTGCTGCCAGCTGCTTCTGGAACCTTGAATTTGATCT 240
QY 241 CAAATGCTACTTCCCTCCAGGAGAGCCCAATAGTCTCCGTCCTCTCAGAGCAGCGTGT 300
Db 241 CAAATGCTACTTCCCTCCAGGAGAGCCCAATAGTCTCCGTCCTCTCAGAGCAGCGTGT 300
QY 301 ATATAAATAAAGGCTAAGATCATATCAGTGATATCATAGTATTAATGGGATTTTC 360
Db 301 ATATAAATAAAGGCTAAGATCATATCAGTGATATCATAGTATTAATGGGATTTTC 360
QY 361 ATATCATAGACAAATTTGCTATCTCCCAAAATTTGCTTATCATCTCCCAAGACAACTCTG 420
Db 361 ATATCATAGACAAATTTGCTATCTCCCAAAATTTGCTTATCATCTCCCAAGACAACTCTG 420
QY 421 GAAGAAATCTGCAAAATCTTACGACTTTGGCAACAAATAGCTACATCAAAATTTAGCA 480
Db 421 GAAGAAATCTGCAAAATCTTACGACTTTGGCAACAAATAGCTACATCAAAATTTAGCA 480
QY 481 ACTTAATACAGACTCAGGTTTCTGAGTGTATCAGGATCATCAGGATCCATCCACCCAGTCA 540
Db 481 ACTTAATACAGACTCAGGTTTCTGAGTGTATCAGGATCATCAGGATCCATCCACCCAGTCA 540
QY 541 CTCTCTTCTGGCCCAACGACCAAGCCCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 CTCTCTTCTGGCCCAACGACCAAGCCCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 TGTTCAACCAAGACAAAGAGCAAGCTGAAGAGTATTTGAAGTTTCAATGATACAG 660
Db 601 TGTTCAACCAAGACAAAGAGCAAGCTGAAGAGTATTTGAAGTTTCAATGATACAG 660
QY 661 ATGCCAAGGTTTACCTGTGATCTTCCACATCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 ATGCCAAGGTTTACCTGTGATCTTCCACATCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 CAGAGCTAGTGAATTTGGAGCTGGCAGGACATCGGTGACCTCTTCTTGAATGCC 780
Db 721 CAGAGCTAGTGAATTTGGAGCTGGCAGGACATCGGTGACCTCTTCTTGAATGCC 780

QY 781 AAACCTGAGAAATTTGTGAGCGGAGCTCTTTGTTTGTGACCTGGTGTGGCTTACGCAATTG 840
Db 781 AAACCTGAGAAATTTGTGAGCGGAGCTCTTTGTTTGTGACCTGGTGTGGCTTACGCAATTG 840
QY 841 ACTGTCTGCTGATTTGATCCCAACCTGGGGGCGCTGTGACACCTTTACTTACTTTCGATG 900
Db 841 ACTGTCTGCTGATTTGATCCCAACCTGGGGGCGCTGTGACACCTTTACTTACTTTCGATG 900
QY 901 CCTCGGGGAGTGTGGGAGCTGTCTAATACTCCACAGTCCCAAGTGGAGTAAACCAA 960
Db 901 CCTCGGGGAGTGTGGGAGCTGTCTAATACTCCACAGTCCCAAGTGGAGTAAACCAA 960
QY 961 AGGTTGTAAGCAGAGTGTCTCTAACAACCTGCTTCAAGAGGAACTCGAAGCGCTGCC 1020
Db 961 AGGTTGTAAGCAGAGTGTCTCTAACAACCTGCTTCAAGAGGAACTCGAAGCGCTGCC 1020
QY 1021 GGGAGCGGTGTCAGCTGCTGATACAGATCCCAAGTCTGCAAGGCTACTTCGGGCGAG 1080
Db 1021 GGGAGCGGTGTCAGCTGCTGATACAGATCCCAAGTCTGCAAGGCTACTTCGGGCGAG 1080
QY 1081 ACTGTGAGGCTGCTGCTGAGGACCAAGATCCCGCTGTAAATAACCGGGGTGTCTGCTTGT 1140
Db 1081 ACTGTGAGGCTGCTGCTGAGGACCAAGATCCCGCTGTAAATAACCGGGGTGTCTGCTTGT 1140
QY 1141 ATCAGTACTCGGCACCGGAGAGTAAATGCAACACCGGCTTCAATGGGACGCGGTGT 1200
Db 1141 ATCAGTACTCGGCACCGGAGAGTAAATGCAACACCGGCTTCAATGGGACGCGGTGT 1200
QY 1201 AGATGTGCTGGCCGGGAGATTCGGGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1201 AGATGTGCTGGCCGGGAGATTCGGGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 GACAGTGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1261 GACAGTGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 GCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Db 1321 GCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 1381 ATGCCACCTGTAAAGAGAAACAACGCTGTGAGTGTAACTGGATTATGAAGGTGACGGA 1440
Db 1381 ATGCCACCTGTAAAGAGAAACAACGCTGTGAGTGTAACTGGATTATGAAGGTGACGGA 1440
QY 1441 TCATATGACAGTGTGATTTCTGCAACACGAGCAACCGGGGCTGTGCAAAAGGTGCGCA 1500
Db 1441 TCATATGACAGTGTGATTTCTGCAACACGAGCAACCGGGGCTGTGCAAAAGGTGCGCA 1500
QY 1501 GATGCTCCCAAGAGGCGAGAGTCTCTGCTGAGCTGCTGAGAGGATCAAAAGGGGACG 1560
Db 1501 GATGCTCCCAAGAGGCGAGAGTCTCTGCTGAGCTGCTGAGAGGATCAAAAGGGGACG 1560
QY 1561 GGCACAGCTGCA CAGAGATAGACCCCTGTGCAAGCGGCTTAAACGAGGCTGTACAGAGC 1620
Db 1561 GGCACAGCTGCA CAGAGATAGACCCCTGTGCAAGCGGCTTAAACGAGGCTGTACAGAGC 1620
QY 1621 AGCCACCTGTAAAGTATGAGGCGGCGGCAAGCA CAAAGTGTGAGTGTAAAGTCACTATG 1680
Db 1621 AGCCACCTGTAAAGTATGAGGCGGCGGCAAGCA CAAAGTGTGAGTGTAAAGTCACTATG 1680
QY 1681 TCGAGATGGGCTGAACCTGTGAGCGGAGAGCTGCCCATTTGACCGCTTACAGGACA 1740
Db 1681 TCGAGATGGGCTGAACCTGTGAGCGGAGAGCTGCCCATTTGACCGCTTACAGGACA 1740
QY 1741 ATGGCAGTGCATGACAGACGCCAAATGTGCGACTCTCCACTTCCAGGATACCACTGTG 1800
Db 1741 ATGGCAGTGCATGACAGACGCCAAATGTGCGACTCTCCACTTCCAGGATACCACTGTG 1800
QY 1801 GGGTGTTCATCTACGCTCCCGCTGCGGCGAGTATAGCTGACCTTTTCAAAAGCCAGAG 1860
Db 1801 GGGTGTTCATCTACGCTCCCGCTGCGGCGAGTATAGCTGACCTTTTCAAAAGCCAGAG 1860

Qy	1861	AGGCCTGTGCGAACGAAGCTGGACCAATGCGAACCTTACAACAGAGTCTCTCTATGCCCCAGA	1920
Db	1861	AGGCCTGTGCGAACGAAGCTGGACCAATGCGAACCTTACAACAGAGTCTCTCTATGCCCCAGA	1920
Qy	1921	AGGCCAAGTACCACTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTCCTACTACCCCA	1980
Db	1921	AGGCCAAGTACCACTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTCCTACTACCCCA	1980
Qy	1981	CAGCCTTCGGCTCCCGAACTGTGGTCTCTGGTGTGGATAGTGGACTATGGACCTTA	2040
Db	1981	CAGCCTTCGGCTCCCGAACTGTGGTCTCTGGTGTGGATAGTGGACTATGGACCTTA	2040
Qy	2041	GACCCAAAGAGTGGAATGTGGATGTCTTCTGCTATCGGATGAAAGATGTGAACCTGCA	2100
Db	2041	GACCCAAAGAGTGGAATGTGGATGTCTTCTGCTATCGGATGAAAGATGTGAACCTGCA	2100
Qy	2101	CCTGCAAGTGGGCTATGTGGGAGATGGCTTCTCATGCACTGGAGACCTGCTGCAGGCTCC	2160
Db	2101	CCTGCAAGTGGGCTATGTGGGAGATGGCTTCTCATGCACTGGAGACCTGCTGCAGGCTCC	2160
Qy	2161	TGATGTCTTTCCTCCTCACTCACAACCTTCCTGACGGAAGTGTCTGGGCTATTTCACACAGCT	2220
Db	2161	TGATGTCTTTCCTCCTCACTCACAACCTTCCTGACGGAAGTGTCTGGGCTATTTCACACAGCT	2220
Qy	2221	CAGCTCGAGCCGTGCATTTCTAGAACCACTGACTGACCTGTCCATCCGCGGCAACCTCT	2280
Db	2221	CAGCTCGAGCCGTGCATTTCTAGAACCACTGACTGACCTGTCCATCCGCGGCAACCTCT	2280
Qy	2281	TTGTGCCACAGAACAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGCGGGACATCGAGC	2340
Db	2281	TTGTGCCACAGAACAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGCGGGACATCGAGC	2340
Qy	2341	ACCACCTCGCCAAATGTCAAGATGTTTTTCTACAAATGACCTTGTCAATGGCACCAACCTGC	2400
Db	2341	ACCACCTCGCCAAATGTCAAGATGTTTTTCTACAAATGACCTTGTCAATGGCACCAACCTGC	2400
Qy	2401	AAACGAGGGTGGGAAGCAAGCTGTCTATCACTGTCCAGCCAGGACCCCACTCCAACCGACGG	2460
Db	2401	AAACGAGGGTGGGAAGCAAGCTGTCTATCACTGTCCAGCCAGGACCCCACTCCAACCGACGG	2460
Qy	2461	AGACCAGGTTGTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTCGCTCCCAATGGGA	2520
Db	2461	AGACCAGGTTGTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTCGCTCCCAATGGGA	2520
Qy	2521	TCATTCAATGTCTTCCAGGCTTTAAAGCACCCCTGCCCCGTGACCTTGACCCACCA	2580
Db	2521	TCATTCAATGTCTTCCAGGCTTTAAAGCACCCCTGCCCCGTGACCTTGACCCACCA	2580
Qy	2581	CTGCTTTGGAGCAGGATCTTTTGGCCATCATCTGTGTGACTGGGGCTGTGTGCTTGG	2640
Db	2581	CTGCTTTGGAGCAGGATCTTTTGGCCATCATCTGTGTGACTGGGGCTGTGTGCTTGG	2640
Qy	2641	CTGCTTACTCTACTTTTCGATAAACCGGAGAACATCGGCTTCCAGCATTTTGTAGTCTGG	2700
Db	2641	CTGCTTACTCTACTTTTCGATAAACCGGAGAACATCGGCTTCCAGCATTTTGTAGTCTGG	2700
Qy	2701	AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTTGAATATCTCGAACCCCT	2760
Db	2701	AAGAGGACATTAATGTTGCAGCTCTTGGCAAGCAGCAGCCTTGAATATCTCGAACCCCT	2760
Qy	2761	TGATATGAGACACAACTCAGCTCCCCAGAACCTTCTCAGACCCCTTCACGGACHTCTG	2820
Db	2761	TGATATGAGACACAACTCAGCTCCCCAGAACCTTCTCAGACCCCTTCACGGACHTCTG	2820
Qy	2821	AAGAACGCGCAGCTTGTAGGGCAATGACCCCTTGAAGACACTGTGAGGGCTTGGACGGAGA	2880
Db	2821	AAGAACGCGCAGCTTGTAGGGCAATGACCCCTTGAAGACACTGTGAGGGCTTGGACGGAGA	2880
Qy	2881	TGCCAGGCATCACTCACTGCGACCTGGGCCCATCAACTGTGAATTTCTCAGCACCAAGTTGCC	2940
Db	2881	TGCCAGGCATCACTCACTGCGACCTGGGCCCATCAACTGTGAATTTCTCAGCACCAAGTTGCC	2940
Qy	2941	TTTTAGGAACGTAAGTCTTTTAAGCACTCAGAAAGCCATACCTCATCTCTCTGGCTGATC	3000

RESULT 3

US-09-842-930A-24
; Sequence 24, Application US/09842930A
; Publication No. US20020197681A1

; FUNCTION NO: 0020
: GENERAL INFORMATION:

APPLICANT: Weigal, Paul

APPLICANT: Weigel, Paul

TITLE OF INVENTION: Identification of Hvaluronan Receptor for Endocytosis

FILE OF INVENTION: IDENTIFICATION
FILE REFERENCE: 5820-603

FILE REFERENCE: 5820.603
CURRENT APPLICATION NUMBER: IIS/09/842.930A

; CURRENT APPLICATION NUMBER: US
 : CURRENT FILING DATE: 2001-04-

;; CURRENT FILING DATE: 2001-04-22
: PRIOR APPLICATION NUMBER: 60/245 320

;; PRIOR APPLICATION NUMBER: 60/

;; PRIOR FILING DATE: 2000-11-02
: PRIOR APPLICATION NUMBER: 60/199 538

;; PRIOR APPLICATION NUMBER: 60/1
PRIOR FILING DATE: 2000-04-25

; PRIOR FILING DATE: 2000-04-21
 ; NUMBER OF SEQ. IN NOS.: 56

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: Pat

; SEQ ID NO 24

; LENGTH: 45

; TYPE: DNA

; ORGANISM: Homo

Query Match 99.2%: Score 3232.6: DB 9: Length 4576:

Query Match

Best Local Similarity 99.7%; Pred. NO: 0;
Matches 2238; Conservative 0; Mismatches 10; Indels 0; Gaps

QY	7	AGCAGGAGCTTCCCAAGAACCCGAAACTTCCCGAGTATTTCTTCCAGTTCGACGAGCATT	66
Db	1328	ATCAGGAGCTTCCCAAGAACCCGAAACTTCCCGAGTATTTCTTCCAGTTCGACGAGCATT	1387
QY	67	TCGTGAAGATCTGGTCGGGCCAGGCCCTTCACTGTTTTGGCACCTTTATCTGCAGCCT	126
Db	1388	TCGTGAAGATCTGGTCGGGCCAGGCCCTTCACTGTTTTGGCACCTTTATCTGCAGCCT	1447
QY	127	TTGATGAGGAAGCTCGGGTTAAAGACTGGACAATAACGGTTTAAATGCCCCAGGTTCCTC	186
Db	1448	TTGATGAGGAAGCTCGGGTTAAAGACTGGACANAATACGGTTTAAATGCCCCAGGTTCCTC	1507
QY	187	GGTACCATGTGGTCCCTGGCCACAGCTGCTTCTGGAAAAACCTGAAATTTGATCTCAATG	246
Db	1508	GGTACCATGTGGTCCCTGGCCACAGCTGCTTCTGGAAAAACCTGAAATTTGATCTCAATG	1567
QY	247	CTACTTCCTCCAAAGGAGAGCCAATAGTCACTCCGTCCTCAGAGCACGGTGTATATAA	306
Db	1568	CTACTTCCTCCAAAGGAGAGCCAATAGTCACTCCGTCCTCAGAGCACGGTGTATATAA	1627
QY	307	ATAATAAGGCTAAGATCATATFCCAGTGATATCATCAGTACTAATPGGATGTTTCATATCA	366
Db	1628	ACATAAGGCTAAGATCATATFCCAGTGATATCATCAGTACTAATPGGATGTTTCATATCA	1687

QY	367	TAGACAAATTTGCTATCTCCCAAAATTTGCTTATCACTCCAAAGACAACTCTCGGAAGAA	426	QY	1447	GCACAGTTGTGGATTTCTGCAAAACAGGACAAACGGGGCTGTGCAAAAGTGGCCAGATGCT	1506
Db	1688	TAGACAAATTTGCTATCTCCCAAAATTTGCTTATCACTCCAAAGACAACTCTCGGAAGAA	1747	Db	2768	GCACAGTTGTGGATTTCTGCAAAACAGGACAAACGGGGCTGTGCAAAAGTGGCCAGATGCT	2827
QY	427	TTCTGCAAAATCTTACGACTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCAACTTAA	486	QY	1507	CCCAGAAGGGCACGAAGGTCTCCTGACAGTCTCCAGAGGGATACAAAGGGGACGGGCACA	1566
Db	1748	TTCTGCAAAATCTTACGACTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCAACTTAA	1807	Db	2828	CCCAGAAGGGCACGAAGGTCTCCTGACAGTCTCCAGAGGGATACAAAGGGGACGGGCACA	2887
QY	487	TACAGGACTCAGGTTTCTCAGGTGTATGACCCGATCCATCCACACCCGAGTCACTCTCT	546	QY	1567	GCTGCAACAGATAGACCCCTGTGACAGACGGCCCTTAAACGAGGGTGTACAGAGCAACCCCA	1626
Db	1808	TACAGGACTCAGGTTTCTCAGGTGTATGACCCGATCCATCCACACCCGAGTCACTCTCT	1867	Db	2888	GCTGCAACAGATAGACCCCTGTGACAGACGGCCCTTAAACGAGGGTGTACAGAGCAACCCCA	2947
QY	547	TCTGGCCACACCGAACCGCTCCATGCCCTTACCTGCTGAACAAACAGGACTTCTGTTTCA	606	QY	1627	CCTGTAAGATGACAGACCGGGCAAGCACAAGGTGTAGTAAAGTCACTATGTGGAG	1686
Db	1868	TCTGGCCACACCGAACCGCTCCATGCCCTTACCTGCTGAACAAACAGGACTTCTGTTTCA	1927	Db	2948	CCTGTAAGATGACAGACCGGGCAAGCACAAGGTGTAGTAAAGTCACTATGTGGAG	3007
QY	607	ACCAAGACAAACAGGACAGCTGAAGAGTATTGAAAGTTTATGATACAGAGTGCCA	666	QY	1687	ATGGGCTGAATCTGTGACCGGAGCAGCTGCCATTTGACCCGCTCTTACAGGACAATGGGC	1746
Db	1928	ACCAAGACAAACAGGACAGCTGAAGAGTATTGAAAGTTTATGATACAGAGTGCCA	1987	Db	3008	ATGGGCTGAATCTGTGACCGGAGCAGCTGCCATTTGACCCGCTCTTACAGGACAATGGGC	3067
QY	667	AGGTTTTAGCTGTGGATCTTCCACATCCACTGCTGGAAGACCCCTGCAAGGTTCAAGC	726	QY	1747	AGTGCCATGACAGACGCAATGTGTGACCTCCACTTCCAGGATACCACTGTTGGGTGT	1806
Db	1988	AGGTTTTAGCTGTGGATCTTCCACATCCACTGCTGGAAGACCCCTGCAAGGTTCAAGC	2047	Db	3068	AGTGCCATGACAGACGCAATGTGTGACCTCCACTTCCAGGATACCACTGTTGGGTGT	3127
QY	727	TGAGTGTGAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTGAATGGCCAAACCT	786	QY	1807	TCCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGCCT	1866
Db	2048	TGAGTGTGAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTGAATGGCCAAACCT	2107	Db	3128	TCCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGAGCCT	3187
QY	787	GCAGAAATGTGACCGGGAGCTCTTGTGACCTGGGTGGGCTACGGCAATGACTGTC	846	QY	1867	GTGCCAAACGAGCTGCGACCATGGCAACCTAACACAGCTCTCTATGCCAGAAGGCCA	1926
Db	2108	GCAGAAATGTGACCGGGAGCTCTTGTGACCTGGGTGGGCTACGGCAATGACTGTC	2167	Db	3188	GTGCCAAACGAGCTGCGACCATGGCAACCTAACACAGCTCTCTATGCCAGAAGGCCA	3247
QY	847	TGCTGATTTGATCCACCTCGGGGGCGCTGTGACACCTTTACTCTTTCGATCGCTCGG	906	QY	1927	AGTACCACCTGTCTCAGCAGCTGCTGGAGACCGGGGGGTGGCTTACCCACAGCCT	1986
Db	2168	TGCTGATTTGATCCACCTCGGGGGCGCTGTGACACCTTTACTCTTTCGATCGCTCGG	2227	Db	3248	AGTACCACCTGTCTCAGCAGCTGCTGGAGACCGGGGGGTGGCTTACCCACAGCCT	3307
QY	907	GGAGGTGTGGAGCTGTGTCAATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG	966	QY	1987	TGGCTCCCAAGACTGTGGTGTGGATAGTGGAGTAGTGGACTATGACACTPAGACCCA	2046
Db	2228	GGAGGTGTGGAGCTGTGTCAATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG	2287	Db	3308	TGGCTCCCAAGACTGTGGTGTGGATAGTGGAGTAGTGGACTATGACACTPAGACCCA	3367
QY	967	TGAAGCAGAGTGTCTCTACACCTGCCCTTCAAGAGAACCTGGAAGGCTGCGGGAGC	1026	QY	2047	ACAAGAGTGAATGTGGGATGTCTTCTGCTATCGGATGAAGATGTGAATGCACCTGCA	2106
Db	2288	TGAAGCAGAGTGTCTCTACACCTGCCCTTCAAGAGAACCTGGAAGGCTGCGGGAGC	2347	Db	3368	ACAAGAGTGAATGTGGGATGTCTTCTGCTATCGGATGAAGATGTGAATGCACCTGCA	3427
QY	1027	GGTGCAGCTGTGTATACAGATCCCGAGTGTGCAAGGCTACTTTCGGGCGAGACTGTC	1086	QY	2107	AGTGGGCTATGTGGGAGATGCTTCTCATGCAAGTGGGACCGGGGGGTGGCTTACCCACAGCCT	2166
Db	2348	GGTGCAGCTGTGTATACAGATCCCGAGTGTGCAAGGCTACTTTCGGGCGAGACTGTC	2407	Db	3428	AGTGGGCTATGTGGGAGATGCTTCTCATGCAAGTGGGACCGGGGGGTGGCTTACCCACAGCCT	3487
QY	1087	AGGCTGCTGCTGGAGGACAGATGCCCCGTGTAAATACCGGGGTGTCTGCCCTTGATCACT	1146	QY	2167	CCTTCCCTCACTCAAACTTCTGACCGAAGTGTGGCCCTATTCCAAACAGCTCAGCTC	2226
Db	2408	AGGCTGCTGCTGGAGGACAGATGCCCCGTGTAAATACCGGGGTGTCTGCCCTTGATCACT	2467	Db	3488	CCTTCCCTCACTCAAACTTCTGACCGAAGTGTGGCCCTATTCCAAACAGCTCAGCTC	3547
QY	1147	ACTCGGCCACCGAGAGTGAATGCAACACCGCTTCAATGGGAGCGGTGTGAGATGT	1206	QY	2227	GAGGCGTGCAATTTCTAGAACACCTGACTGCTTCCATCCCGGCAACCTCTTTGTGC	2286
Db	2468	ACTCGGCCACCGAGAGTGAATGCAACACCGCTTCAATGGGAGCGGTGTGAGATGT	2527	Db	3548	GAGGCGTGCAATTTCTAGAACACCTGACTGCTTCCATCCCGGCAACCTCTTTGTGC	3607
QY	1207	GCTGGCCGGGGAGATTCGGGCTGATTTGTCTGCCCTGTGGCTGCTCAGACCAACGACACT	1266	QY	2287	CACAGAACAGTGGGCTGGGGAGATGAGACCTGTCTGGGGGGGACATCGAGCACCAAC	2346
Db	2528	GCTGGCCGGGGAGATTTGGGCTGATTTGTCTGCCCTGTGGCTGCTCAGACCAACGACACT	2587	Db	3608	CACAGAACAGTGGGCTGGGGAGATGAGACCTGTCTGGGGGGGACATCGAGCACCAAC	3667
QY	1267	GGGATGATGGATACAGGGCTCCGGCAGTGTCTGTGAACACGGGGTGGACAGGCCCT	1326	QY	2347	TCGCCAATGTGAGCATGTTTTTCTACATGACCTTTGCAATGGCCACCACTTCGCAACCA	2406
Db	2588	GGGATGATGGATACAGGGCTCCGGCAGTGTCTGTGAACACGGGGTGGACAGGCCCT	2647	Db	3668	TCGCCAATGTGAGCATGTTTTTCTACATGACCTTTGCAATGGCCACCACTTCGCAACCA	3727
QY	1327	CGTGTGACACTCAGGCAAGTTTTGCTGACGTGTGACCGCTCTTGTGCTCTCATGCCA	1386	QY	2407	GGTGGGAGCAGCTGCTCATCTGCGCAGCAGGACCCACTCCAAACGAGGAGACCA	2466
Db	2648	CGTGTGACACTCAGGCAAGTTTTGCTGACGTGTGACCGCTCTTGTGCTCTCATGCCA	2707	Db	3728	GGCTGGGAGCAAGCTGCTCATCTGCGCAGCAGGACCCACTCCAAACGAGGAGACCA	3787
QY	1387	CCTGTAAGAGAACACAGTGTGAGTGAACCTGGATTAAGAGTGACGGAATCACAT	1446	QY	2467	GGTTTTGTTGATGGAGAGCCATTCTGACAGTGGGACATCTTTGCTCCCAATGGATCATTC	2526
Db	2708	CCTGTAAGAGAACACAGTGTGAGTGAACCTGGATTAAGAGTGACGGAATCACAT	2767	Db	3788	GGTTTTGTTGATGGAGAGCCATTCTGACAGTGGGACATCTTTGCTCCCAATGGATCATTC	3847
				QY	2527	ATGTCATTTCCAGGGCTTTAAAGCAACCCCTGCCCCGTGACCTTGACCTTGACCCACACTGGCT	2586

Db 4574 CCCTAGCCAGTCTGCACAGGAAGTGTGCACAAATAAGGTTTATGGAAACAGAAACAA 4633
Qy 3247 AAAAAAAAA 3254
Db 4634 AGTCAACA 4641

RESULT 5

US-10-028-248A-1
; Sequence 1, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennnda
; APPLICANT: Zethusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163)..(168)
; OTHER INFORMATION: wherein n is A, C, G, or T
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1279)..(1284)
; OTHER INFORMATION: wherein n is A, C, G, or T

US-10-028-248A-1

Query Match 92.0%; Score 2999.8; DB 16; Length 8444;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 3128; Conservative 0; Mismatches 18; Indels 108; Gaps 1;

Qy 7 AGCAGGAGCTTCCCAAGAACCCGAAACCTTCCCAAGTATTTCTTCAGTGTGAGGAGCAATT 66
Db 5279 ATCAGGAGCTTCCCAAGAACCCGAAACCTTCCCAAGTATTTCTTCAGTGTGAGGAGCAATT 5338
Qy 67 TCSTGAAAGATCTGCTGGCCCGAGCCCTTCACTGTTTTTGGACCTTTATCTGCAGGCT 126
Db 5339 TCGTAAAGATCTGCTGGCCCGAGCCCTTCACTGTTTTTGGACCTTTATCTGCAGGCT 5398
Qy 127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCAGGTTCTTTC 186
Db 5399 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCAGGTTCTTTC 5458
Qy 187 GGTACCATGTGCTGCGCTGCCACACAGCTGCTTCTGGAACCTGAAATGATCTCAATG 246
Db 5459 GGTACCATGTGCTGCGCTGCCACACAGCTGCTTCTGGAACCTGAAATGATCTCAATG 5518
Qy 247 CTACTTCCCTCCAAAGGAGAGCAATAGTCATCTCGTCTCTCAGAGCACGGTGTATATAA 306
Db 5519 CTACTTCCCTCCAAAGGAGAGCAATAGTCATCTCGTCTCTCAGAGCACGGTGTATATAA 5578
Qy 307 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTTGTTCAATCA 366
Db 5579 ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTTGTTCAATCA 5638
Qy 367 TAGACAAATTCGTATCTCCAAAATTTGCTTATCATCTCCAAAGACAACTCTGGAAGAA 426
Db 5639 TAGACAAATTCGTATCTCCAAAATTTGCTTATCATCTCCAAAGACAACTCTGGAAGAA 5698
Qy 427 TTCTGCAAAATCTTACGACTTTGGCAACAAACATAGGTACATCAATTTTAGCAACTTAA 486
Db 5699 TTCTGCAAAATCTTACGACTTTGGCAACAAACATAGGTACATCAATTTTAGCAACTTAA 5758
Qy 487 TACAGGACTCAGGTTTGTGAGTGTATCATCCGATCCCATCCACACCCAGTCACTCTCT 546
Db 5759 TACAGGACTCAGGTTTGTGAGTGTATCATCCGATCCCATCCACACCCAGTCACTCTCT 5818
Qy 547 TCTGCCACCGAACCAAGCCCTCGATGCCCTTACCTGTCTGAACACAGAGACTTCTGTGTA 606
Db 5819 TCTGCCACCGAACCAAGCCCTCGATGCCCTTACCTGTCTGAACACAGAGACTTCTGTGTA 5878
Qy 607 ACCAAGACAAACAGGACCAAGCTGGAAGAGTATTTGAAGTTTCATGTGATACGAGATGCA 666
Db 5879 ACCAAGACAAACAGGACCAAGCTGGAAGAGTATTTGAAGTTTCATGTGATACGAGATGCA 5938
Qy 667 AGGTTTATAGTGTGGATCTTCCACATCCACTGCTGGAAGACCTGCAAGGTTGAGAGC 726
Db 5939 AGGTTTATAGTGTGGATCTTCCACATCCACTGCTGGAAGACCTGCAAGGTTGAGAGC 5998
Qy 727 TGAGTGTGAATGTGGAGCTGCGAGGACATCGGTGACCTCTTTCTGAATGCCAAACCT 786
Db 5999 TGAGTGTGAATGTGGAGCTGCGAGGACATCGGTGACCTCTTTCTGAATGCCAAACCT 6058
Qy 787 GCAGAAATGTGCAGCGGAGCTCTTTGTTGACCTGGGTGTGCCCTACGGCATTGACTGTC 846
Db 6059 GCAGAAATGTGCAGCGGAGCTCTTTGTTGACCTGGGTGTGCCCTACGGCATTGACTGTC 6118
Qy 847 TGCTGATTGATCCCAACCTTGGGGGCGCGCTGTGACACCTTTTACTTTCGATGCTCGG 906
Db 6119 TGCTGATTGATCCCAACCTTGGGGGCGCGCTGTGACACCTTTTACTTTCGATGCTCGG 6178
Qy 907 GGGAGTGTGGAGCTGTGTCAATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTGT 966
Db 6179 GGGAGTGTGGAGCTGTGTCAATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTGT 6238
Qy 967 TGAACGAGAAGTGTCTTACACCTGCGCTTCAAGAGGAACCTGGAAGGCTCCCGGAGC 1026
Db 6239 TGAACGAGAAGTGTCTTACACCTGCGCTTCAAGAGGAACCTGGAAGGCTCCCGGAGC 6298
Qy 1027 GGTGACGCTGTGTATACAGATCCCAAGGTGTGCAAGGGCTACTTTCGGGCGAGACTGTC 1086
Db 6299 GGTGACGCTGTGTATACAGATCCCAAGGTGTGCAAGGGCTACTTTCGGGCGAGACTGTC 6358


```
Db      8411 AGTCAAAATAAAAA 8424

RESULT 6
US-10-107-782-1
; Sequence 1, Application US/10107782
; Publication No. US20040618970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malvankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glenda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zethusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 1
; LENGTH: 8444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8025)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (163)..(163)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (164)..(164)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (165)..(165)
; OTHER INFORMATION: Wherein n may be a, c, g or t

;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (166)..(166)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (167)..(167)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (168)..(168)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1279)..(1279)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1280)..(1280)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1281)..(1281)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1282)..(1282)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1283)..(1283)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1284)..(1284)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (163)..(165)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (166)..(168)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1279)..(1281)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1282)..(1284)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; US-10-107-782-1
Query Match      92.0%; Score 2999.8; DB 16; Length 8444;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 3128; Conservative 0; Mismatches 18; Indels 108; Gaps 1;

QY      7  AGCAGGAGCTCCCAAGAACCCGAAACTCCAGTATTTCTCCAGTTGCAGGAGCATT 66
Db      5279 ATCAGGAGCTTCCCAAGAACCCGAAACTTCCAGTATTTCTCCAGTTGCAGGAGCATT 5338
QY      67  TCGTGAAGAATCTGTCGGCCCGAGCCCTTCACTGTTTTCACCTTTATCTGCAGCCT 126
Db      5339 TCGTGAAGAATCTGTCGGCCCGAGCCCTTCACTGTTTTCACCTTTATCTGCAGCCT 5398
QY      127 TTGATGAGGAGCTCGGTTAAAGACTGGACAAATACGTTTAAATGCCAGGTTCTTC 186
Db      5399 TTGATGAGGAGCTCGGTTAAAGACTGGACAAATACGTTTAAATGCCAGGTTCTTC 5458
QY      187 GGTACCATGTGTCGCTCGCCCTGCCACAGCTGCTTCTGAAAAACCTGAAATGATCTCAAAATG 246
Db      5459 GGTACCATGTGTCGCTCGCCCTGCCACAGCTGCTTCTGAAAAACCTGAAATGATCTCAAAATG 5518
```


QY	247	CTACTTCCTCCAAAGGAGGCAATAGTATCATCTCCGCTCTCTCAGAGCAAGTGTATATA	306
Db	5519	CTACTTCCTCCAAAGGAGGCAATAGTATCATCTCCGCTCTCTCAGAGCAAGTGTATATA	5578
QY	307	ATAATAAGGCTAAGATCATATCCAGTGAATATCATAGTACTAATGGATTGTTCAATCA	366
Db	5579	ATAATAAGGCTAAGATCATATCCAGTGAATATCATAGTACTAATGGATTGTTCAATCA	5638
QY	367	TAGACAAATGCTATCTCCAAAATTTGCTTATCACTCCCAAAGCAACTCTGGAAGAA	426
Db	5639	TAGACAAATGCTATCTCCAAAATTTGCTTATCACTCCCAAAGCAACTCTGGAAGAA	5698
QY	427	TTCTGCAAAATCTTACGACTTTGGCAACAAATGCTACATCAAAATTTAGCAACTTAA	486
Db	5699	TTCTGCAAAATCTTACGACTTTGGCAACAAATGCTACATCAAAATTTAGCAACTTAA	5758
QY	487	TACAGACTCAGGTTTGTCTGAGTGTATCAACCGATCCCATCCACCCCAAGTCACTCTCT	546
Db	5759	TACAGACTCAGGTTTGTCTGAGTGTATCAACCGATCCCATCCACCCCAAGTCACTCTCT	5818
QY	547	TCTGGCCCAACGACCAAGCCCTCCATGCCCTACTGCTGAAACAAAGGACTTCTGTTCA	606
Db	5819	TCTGGCCCAACGACCAAGCCCTCCATGCCCTACTGCTGAAACAAAGGACTTCTGTTCA	5878
QY	607	ACCAAGACAACAAAGGACAAGCTGAAGAGTATTTGAAGTTTCATGTGATACGAGATGCCA	666
Db	5879	ACCAAGACAACAAAGGACAAGCTGAAGAGTATTTGAAGTTTCATGTGATACGAGATGCCA	5938
QY	667	AGGTTTATGCTGTGGATCTTCCOACATCCATGCTGGAAGACCTTGCAGAGTTTCAGAGC	726
Db	5939	AGGTTTATGCTGTGGATCTTCCOACATCCATGCTGGAAGACCTTGCAGAGTTTCAGAGC	5998
QY	727	TGAGTGTGAATCTGAGCTGCGCAGGACATCGGTGACCTTCTGTAATGGCAACCT	786
Db	5999	TGAGTGTGAATCTGAGCTGCGCAGGACATCGGTGACCTTCTGTAATGGCAACCT	6058
QY	787	GCAGAAATCTGACGCGGAGCTCTTGTGTTGACCTGGGTGTGGCTACGGCATTTGACTGTC	846
Db	6059	GCAGAAATCTGACGCGGAGCTCTTGTGTTGACCTGGGTGTGGCTACGGCATTTGACTGTC	6118
QY	847	TGCTGATTTGATCCCACTCGGGGGCCGCTGTGACACCTTTACTATCTTTCGATGCTCGG	906
Db	6119	TGCTGATTTGATCCCACTCGGGGGCCGCTGTGACACCTTTACTATCTTTCGATGCTCGG	6178
QY	907	GGGAGTGTGGAGCTGTGTCATATCTCCAGCTGCCAGGTGGAGTAACCAAGGGTG	966
Db	6179	GGGAGTGTGGAGCTGTGTCATATCTCCAGCTGCCAGGTGGAGTAACCAAGGGTG	6238
QY	967	TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGAGC	1026
Db	6239	TGAAGCAGAAGTGTCTCTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGAGC	6298
QY	1027	GCTGACGCTGTGTATACAGATCCCGAGTGTGCAAGGGCTACTTTGGGGGAGACTGTTC	1086
Db	6299	GCTGACGCTGTGTATACAGATCCCGAGTGTGCAAGGGCTACTTTGGGGGAGACTGTTC	6358
QY	1087	AGGCTTCCCTCGAGGACAGATGCCCGTGTAAATAACCGGGGTGTCTGCTTGTATCAGT	1146
Db	6359	AGGCTTCCCTCGAGGACAGATGCCCGTGTAAATAACCGGGGTGTCTGCTTGTATCAGT	6418
QY	1147	ACTCGGCCACCGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT	1206
Db	6419	ACTCGGCCACCGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT	6478
QY	1207	GCTGGCCGGGAGATTCGGGCTGATTTGCTGCCCTGTGGCTGTCTGACACCGACAGT	1266
Db	6479	GCTGGCCGGGAGATTCGGGCTGATTTGCTGCCCTGTGGCTGTCTGACACCGACAGT	6538
QY	1267	GGGATGATGGCATCAGGGCTCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCT	1326
Db	6539	GGGATGATGGCATCAGGGCTCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCT	6598
QY	1327	CGTGTGACACTCAGGCAGTTTTTGCCTCGAGTGTGTAGCCTCTTGTCTGCTCATGCCA	1386
Db	6599	CGTGTGACACTCAGGCAGTTTTTGCCTCGAGTGTGTAGCCTCTTGTCTGCTCATGCCA	6658
QY	1387	CCTGTAAGGAGAACACAGCTGTAGTGTAACTCGATTATGAAGGTGACGGAATCACAT	1446
Db	6659	CCTGTAAGGAGAACACAGCTGTAGTGTAACTCGATTATGAAGGTGACGGAATCACAT	6718
QY	1447	GACAGTTGTGATTTCTGAAACACAGCAACCGGGGCTGTGCAAAAGTGGCCAGATGCT	1506
Db	6719	GACAGTTGTGATTTCTGAAACACAGCAACCGGGGCTGTGCAAAAGTGGCCAGATGCT	6778
QY	1507	CCAGAAAGGCGACGAAGTCTCTGTCAGCTGCCAGAGGAGATACAAAGGGGACGGGACA	1566
Db	6779	CCAGAAAGGCGACGAAGTCTCTGTCAGCTGCCAGAGGAGATACAAAGGGGACGGGACA	6838
QY	1567	GCTGCACAGAGATAGACCCCTGTGCAGACGGCCCTTAAACGGAGGCTGTACAGCACGCCA	1626
Db	6839	GCTGCACAGAGATAGACCCCTGTGCAGACGGCCCTTAAACGGAGGCTGTACAGCACGCCA	6898
QY	1627	CCTGTAAGATGACAGGCCCGGGAGACCAAGTGTGAGTGTAAAGTCACTATGTGGAG	1686
Db	6899	CCTGTAAGATGACAGGCCCGGGAGACCAAGTGTGAGTGTAAAGTCACTATGTGGAG	6958
QY	1687	ATGGGCTGAATCTGAGCGGAGCAGCTGCCCATTTGACCGCTTACAGGACAATGGGC	1746
Db	6959	ATGGGCTGAATCTGAGCGGAGCAGCTGCCCATTTGACCGCTTACAGGACAATGGGC	7018
QY	1747	AGTGCATGACAGACGCAAAATGTGCGACCTCCACTTCCAGGATACACTGTTGGGGTGT	1806
Db	7019	AGTGCATGACAGACGCAAAATGTGCGACCTCCACTTCCAGGATACACTGTTGGGGTGT	7078
QY	1807	TCCATCTAGCTCCCACTGGGCCAGTAAAGTGTGAGTGTAAAGTCACTATGTGGAGCCT	1866
Db	7079	TCCATCTAGCTCCCACTGGGCCAGTAAAGTGTGAGTGTAAAGTCACTATGTGGAGCCT	7138
QY	1867	GTGCCAAAGAGCTGCGACCATGGCAACCTCAACCCAGCTCTCTATGCCAGAGGCCA	1926
Db	7139	GTGCCAAAGAGCTGCGACCATGGCAACCTCAACCCAGCTCTCTATGCCAGAGGCCA	7198
QY	1927	AGTACCACTGTGCTCAGCAGCTGGCTGGAGACCGGGGGGTGCTTACCCCAACAGCCT	1986
Db	7199	AGTACCACTGTGCTCAGCAGCTGGCTGGAGACCGGGGGGTGCTTACCCCAACAGCCT	7258
QY	1987	TGCGCTCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2046
Db	7259	TGCGCTCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	7318
QY	2047	ACAAGAGTGAATGTGGGATGTCTTCTGCTATCGGATGAAGATGTGAACCTGCACCTGCA	2106
Db	7319	ACAAGAGTGAATGTGGGATGTCTTCTGCTATCGGATGNA-----	7358
QY	2107	AGGTGGGCTATGTGGAGATGGCTTCTCATGCTGAGTGGGAACTGTGCTGAGTCTGATGT	2166
Db	7359	-----	7358
QY	2167	CCTTCCCTCACTCACAACCTTCTGACGGAAGTGTGCTTCTTCCACAGCTCAGGTC	2226
Db	7359	-----GGAAGTGTGCTGCTTCTTCCACAGCTCAGGTC	7390
QY	2227	GAGGCGGTGCAATTTCTAGAACACCTGTGCTGCTTCCATCCGCGGACCCCTCTTTGTGC	2286
Db	7391	GAGGCGGTGCAATTTCTAGAACACCTGTGCTGCTTCCATCCGCGGACCCCTCTTTGTGC	7450
QY	2287	CACAGAACAGTGGGCTGGGGAGATGAGACTTGTCTGGGGGGGAGATCGAGCACACC	2346
Db	7451	CACAGAACAGTGGGCTGGGGAGATGAGACTTGTCTGGGGGGGAGATCGAGCACACC	7510
QY	2347	TGCGCAATGTGAGCATGTTTTTCTCAATGACTTGTCAATGGCACCCCTGCAAAACCA	2406
Db	7511	TGCGCAATGTGAGCATGTTTTTCTCAATGACTTGTCAATGGCACCCCTGCAAAACCA	7570
QY	2407	GGGTGGGAAGCAAGCTGCTCATCTGCTGCGACGCAACCCACTCCAAACGACGAGACCA	2466

Db 7571 GGCTGGGAAGCAAGTGTCTCATCTGCGCAGCAGGACCCACTCAACCGGTACAAAGTA 7630
QY 2467 GGTGTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCCTCCAAATGGGATCATTC 2526
Db 7631 GGTGTGTTGATGGAAGAGCCATTCTGCAGTGGGACATCTTTGCCCTCCAAATGGGATCATTC 7690
QY 2527 ATGTCAATTTCCAGGCGCTTTAAAGACACCCCTTGCCCGGTGACCTTGACCCACATGGCT 2586
Db 7691 ATGTCAATTTCCAGGCGCTTTAAAGACACCCCTTGCCCGGTGACCTTGACCCACATGGCT 7750
QY 2587 TGGGAGCAGGATCTTCTTTGCGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2646
Db 7751 TGGGAGCAGGATCTTCTTTGCGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7810
QY 2647 ACTCTACTTTCCGATAAACCCGGAACAATCGGCTTCAGCATTTTGAAGCGAAGAGG 2706
Db 7811 ACTCTACTTTCCGATAAACCCGGAACAATCGGCTTCAGCATTTTGAAGCGAAGAGG 7870
QY 2707 ACATTAATGTTTCAGCTCTTGGCAAGCAGCAGCTTGAGAATATCTGAAACCCCTTTGTATG 2766
Db 7871 ACATTAATGTTTCAGCTCTTGGCAAGCAGCAGCTTGAGAATATCTGAAACCCCTTTGTATG 7930
QY 2767 AGAGCACACCTCAGCTCCCGCAGAACCTTCTAGACCCCTTCAGGACTCTGAAGAC 2826
Db 7931 AGAGCACACCTCAGCTCCCGCAGAACCTTCTAGACCCCTTCAGGACTCTGAAGAC 7990
QY 2827 GGCAGCTTGAGGCGAATACCCCTTTGAGGACACTGTGAGGGGCTTGAGCGGAGATGCCAG 2886
Db 7991 GGCAGCTTGAGGCGAATACCCCTTTGAGGACACTGTGAGGGGCTTGAGCGGAGATGCCAG 8050
QY 2887 CCATCACTCAGTGCACCTGGGCCATCACTGTGAATTTCTAGACCCAGTGTGCCCTTTAG 2946
Db 8051 CCATCACTCAGTGCACCTGGGCCATCACTGTGAATTTCTAGACCCAGTGTGCCCTTTAG 8110
QY 2947 GAACGTAAAGTCTTTAAGCACTCAGAGCCACTCATCTCTGCTGCTGATCTGGGG 3006
Db 8111 GAACGTAAAGTCTTTAAGCACTCAGAGCCACTCATCTCTGCTGCTGATCTGGGG 8170
QY 3007 TTGTTTCTGTGGTGAGAGATGTGTTGCTGCTGCCACCCAGTACAGCTTCCCTCTGAC 3066
Db 8171 TTGTTTCTGTGGTGAGAGATGTGTTGCTGCTGCCACCCAGTACAGCTTCCCTCTGAC 8230
QY 3067 CTTTGGCTCTTCTTCTTGTGACTCTTCAGTGGCAGCTGCTCCATTTGCTGCTTACATG 3126
Db 8231 CTTTGGCTCTTCTTCTTGTGACTCTTCAGTGGCAGCTGCTCCATTTGCTGCTTACATG 8290
QY 3127 ATGGGTAAGTGTGATCTTCTTCTTGTGACTCTTCAGTGGCAGCTGCTCCATTTGCTGCTTACATG 3186
Db 8291 ATGGGTAAGTGTGATCTTCTTCTTGTGACTCTTCAGTGGCAGCTGCTCCATTTGCTGCTTACATG 8350
QY 3187 CCCTAGCCAGTGCCTGACACAGGAACCTGTGCACAAATAAGGTTTATGGAAACAGAAAAA 3246
Db 8351 CCCTAGCCAGTGCCTGACACAGGAACCTGTGCACAAATAAGGTTTATGGAAACAGAAAAA 8410
QY 3247 AAAAAAATAAAAAA 3260
Db 8411 AGTCAAAAAAATAAAAAA 8424

RESULT 7

US-10-028-248A-3
; Sequence 3, Application US/10028248A
; Publication NO. US20030235882A1
; GENERAL INFORMATION:

; APPLICANT: Shimketa, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha

; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 8495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131)..(1316)
; OTHER INFORMATION: Wherein n is A, C, G, or T
US-10-028-248A-3

Query Match 91.9%; Score 2995.6; DB 16; Length 8495;
Best Local Similarity 96.1%; Pred No. 0;
Matches 3137; Conservative 0; Mismatches 10; Indels 116; Gaps 2;

QY 7 AGCAGGAGCTTCCCAAGAACCCGAAACTTCCAGTATTTCTTCCAGTTGCGAGGCAATT 66
Db 5311 ATCAGGAGCTTCCCAAGAACCCGAAACTTCCAGTATTTCTTCCAGTTGCGAGGCAATT 5370
QY 67 TCGTGAAGATCTGCTCGGCCCGAGCCCTTCACTGTTTTCACCTTTATTCGAGCT 126
Db 5371 TCGTGAAGATCTGCTCGGCCCGAGCCCTTCACTGTTTTCACCTTTATTCGAGCT 5430
QY 127 TTGATGAGCAAGCTCGGGTTAAAGACTGGGCAAAATACGGTTTAAATGCCAGGTTCTTC 186
Db 5431 TTGATGAGCAAGCTCGGGTTAAAGACTGGGCAAAATACGGTTTAAATGCCAGGTTCTTC 5490
QY 187 GGTACCATGTGCTCGCTGCCCGCAGCTGCTTCTTGAAAAACCTGAAATTTGATCTCAAAATG 246
Db 5491 GGTACCATGTGCTCGCTGCCCGCAGCTGCTTCTTGAAAAACCTGAAATTTGATCTCAAAATG 5550
QY 247 CTACTTCCCTCCAGGAGAGCCCAATAGTCATCTCGTCTCTCAGNAGCAGGTTGATATAA 306
Db 5551 CTACTTCCCTCCAGGAGAGCCCAATAGTCATCTCGTCTCTCAGNAGCAGGTTGATATAA 5610
QY 307 ATAATAAGGCTAAGATCATATTCAGTGTATATCATCTAGTAAATGGGATTGTTTCATATCA 366
Db 5611 ACAATAAGGCTAAGATCATATTCAGTGTATATCATCTAGTAAATGGGATTGTTTCATATCA 5670

QY	367	TAGACAAATTGCTATCTCCAAAAATTTGCTTATCACTCCAAAGACAACTCTGGAGAA	426	Db	6751	GAATCACATGSCAGTTGTGGATTTCTGCAAAACAGACAAACGCGGCTGTGCAAGGTGG	6810
Db	5671	TAGACAAATTGCTATCTCCAAAAATTTGCTTATCACTCCAAAGACAACTCTGGAGAA	5730	QY	1498	CCAGATGCTCCCAAGAGGGGACGAAGGTCTCTGAGAGTCTGCGAGAGGTATCAAGAGGG	1557
QY	427	TTCTGCAAAATCTTACGATTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCACTTAA	486	Db	6811	CCAGATGCTCCCAAGAGGGGACGAAGGTCTCTGAGAGTCTGCGAGAGGTATCAAGAGGG	6870
Db	5731	TTCTGCAAAATCTTACGATTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCACTTAA	5790	QY	1558	ACGGGCACAGCTGCACAGAGATAGACCCCTGTGACAGCGGCTTAAACGGAGGGTGTACAG	1617
QY	487	TACAGCACTCAGGTTTCTGAGTGATCATCACCGATCCCATCCACACCCAGTCACTCTCT	546	Db	6871	ACGGGCACAGCTGCACAGAGATAGACCCCTGTGACAGCGGCTTAAACGGAGGGTGTACAG	6930
Db	5791	TACAGCACTCAGGTTTCTGAGTGATCATCACCGATCCCATCCACACCCAGTCACTCTCT	5850	QY	1618	AGCACGCCACCTGTAAAGATGACAGGCGCGGCAAGCACAAAGTGTAGTGTAAAGTCACT	1677
QY	547	TCGTGGCCCAACGACCAAA-----GCCCTCCATGCCCTTACCTGCTGAACACAGGACT	597	Db	6931	AGCACGCCACCTGTAAAGATGACAGGCGCGGCAAGCACAAAGTGTAGTGTAAAGTCACT	6990
Db	5851	TCGTGGCCCAACGACCAACGACCTCCATGCGCTTCAATGCCCTTACCTGCTGAACACAGGACT	5910	QY	1678	ATGTCCGAGATGGGCTGAACTGTGAGCCGAGGAGCTGCCCATTTGACCCGTGCTTACAG	1737
QY	598	TCCTGTTCAACCAAGACAACAAGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATAC	657	Db	6991	ATGTCCGAGATGGGCTGAACTGTGAGCCGAGGAGCTGCCCATTTGACCCGTGCTTACAG	7050
Db	5911	TCCTGTTCAACCAAGACAACAAGACAAGCTGAAGGAGTATTTGAAGTTTCATGTGATAC	5970	QY	1738	ACAAATGGGCAGTGCATGACAGCGCCAAATGTGTCCACCTCCACTTTCCAGGATACCACTG	1797
QY	658	GAGATCCCAAGTTTGTAGCTGTGGATCTTCCACATCCACTGCTCGAGAGCCCTCAAG	717	Db	7051	ACAAATGGGCAGTGCATGACAGCGCCAAATGTGTCCACCTCCACTTTCCAGGATACCACTG	7110
Db	5971	GAGATCCCAAGTTTGTAGCTGTGGATCTTCCACATCCACTGCTCGAGAGCCCTCAAG	6030	QY	1798	TTGGGCTGTTCATCTTACCTCTCCACTCGGCGCAGTATTAAGCTGACCTTTGACAAAGCCA	1857
QY	718	GTTTCAGAGCTGAGTGTAAATGTGAGCTGGCAGGAGACATCGGTGACCTCTTTCTGAATG	777	Db	7111	TTGGGCTGTTCATCTTACCTCTCCACTCGGCGCAGTATTAAGCTGACCTTTGACAAAGCCA	7170
Db	6031	GTTTCAGAGCTGAGTGTAAATGTGAGCTGGCAGGAGACATCGGTGACCTCTTTCTGAATG	6090	QY	1858	GAGAGCCCTGTGCGCAACGAGCTGCGACCATGGCAACCTTACACAGCTCTCTTATGCC	1917
QY	778	GCCAAACCTGCAGAAATGTGAGCGGAGCTCTTTGTTGACCTTGGGTGTGGCCCTACGGCA	837	Db	7171	GAGAGCCCTGTGCGCAACGAGCTGCGACCATGGCAACCTTACACAGCTCTCTTATGCC	7230
Db	6091	GCCAAACCTGCAGAAATGTGAGCGGAGCTCTTTGTTGACCTTGGGTGTGGCCCTACGGCA	6150	QY	1918	AGAAGCCCAAGTACCACTGTGCTCAGCAGGCTGGCTGGAGACCGGCGGCTTGCCTACC	1977
QY	838	TTGACTGTCTGCTGATTTGATCCACCTCGGGGCGGCTGTGACACTTTTACTACTTTCG	897	Db	7231	AGNAGGCCAAGTACCACTGTGCTCAGCAGGCTGGCTGGAGACCGGCGGCTTGCCTACC	7290
Db	6151	TTGACTGTCTGCTGATTTGATCCACCTCGGGGCGGCTGTGACACTTTTACTACTTTCG	6210	QY	1978	CCAAGCCCTTCGCTCCCAAGAACTGTGCTCTGGTGTGGTGTGGATAGTGGACTATGGAC	2037
QY	898	ATGCTTCGGGGAGTGTGGAGCTGTGTCAATACTCCAGCTGCCCAAGGTGAGTAAAC	957	Db	7291	CCAAGCCCTTCGCTCCCAAGAACTGTGCTCTGGTGTGGTGTGGATAGTGGACTATGGAC	7350
Db	6211	ATGCTTCGGGGAGTGTGGAGCTGTGTCAATACTCCAGCTGCCCAAGGTGAGTAAAC	6270	QY	2038	CTAGACCCCAAGAGTGAATGTGGATGTCTTCTGCTATCGGATGGAAGATGTGACT	2097
QY	958	CAAGGGTGTGAAGCAAGTGTCTCTCAACCTGCGCTTCAAGAGAACCTGGAAGGCT	1017	Db	7351	CTAGACCCCAAGAGTGAATGTGGATGTCTTCTGCTATCGGATGGAAGATGTGACT	7400
Db	6271	CAAGGGTGTGAAGCAAGTGTCTCTCAACCTGCGCTTCAAGAGAACCTGGAAGGCT	6330	QY	2098	GCACCTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCACTGGGAAACCTGCTGAGG	2157
QY	1018	GCGGAGCGGTGACGCTGTGATACAGATCCCAAGTGTCTCAAGGGCTACTTCGGGC	1077	Db	7401	-----	7400
Db	6331	GCGGAGCGGTGACGCTGTGATACAGATCCCAAGTGTCTCAAGGGCTACTTCGGGC	6390	QY	2158	TCTGATGTCTTCCCTCACTCAAACTTCTGACGGAAGTGTGGCTATTTCCAACA	2217
QY	1078	GAGACTGTGAGGCTGCGCTGGAGGACAGATGCCCGGTGTAAATACCGGGGTGTCTGCC	1137	Db	7401	-----GGAAGTGTGGCTATTTCCAACA	7423
Db	6391	GAGACTGTGAGGCTGCGCTGGAGGACAGATGCCCGGTGTAAATACCGGGGTGTCTGCC	6450	QY	2218	GCTCAGCTCGAGGCGTGCATTTCTAGAACACTGACTGACTGCTCCATCCGCGGACCC	2277
QY	1138	TTGATCAGTACTCGGCCACCGGAGAGTGAATGCAACACCGGCTTCAATGGACGGCGT	1197	Db	7424	GCTCAGCTCGAGGCGTGCATTTCTAGAACACTGACTGACTGCTCCATCCGCGGACCC	7483
Db	6451	TTGATCAGTACTCGGCCACCGGAGAGTGAATGCAACACCGGCTTCAATGGACGGCGT	6510	QY	2278	TCCTTTGTGCAACAGACAGTGGGCTGGGAGAAATGAGACCTTGTCTGGCGGGGACATCG	2337
QY	1198	GTGAGATGTCTGCGCGGGAGATTCGGGCTCAATGTCTGCGCTGTGGCTGTCTCAGACC	1257	Db	7484	TCCTTTGTGCAACAGACAGTGGGCTGGGAGAAATGAGACCTTGTCTGGCGGGGACATCG	7543
Db	6511	GTGAGATGTCTGCGCGGGAGATTTGGGCTCAATGTCTGCGCTGTGGCTGTCTCAGACC	6570	QY	2338	AGCACCACTCTGCGCAATGTGAGATGTTTTTTCTAATGACCTTGTCAATGGACCAACCC	2397
QY	1258	ACGGAAGTGCATGATGGAATCAACGGGCTCGGGGAGTGCCTCTGTGAAACGGGTTGA	1317	Db	7544	AGCACCACTCTGCGCAATGTGAGATGTTTTTTCTAATGACCTTGTCAATGGACCAACCC	7603
Db	6571	ACGGAAGTGCATGATGGAATCAACGGGCTCGGGGAGTGCCTCTGTGAAACGGGTTGA	6630	QY	2398	TGCAAAACGAGGTTGGGAAGCAAGTCTCATCACTGCCAGCAGGACCCACTTCCAACCGA	2457
QY	1318	CAGGCCCCCTGCTGACACTCAGGAGTTTTCCTGCTCAGTGATGACCTCTTGTCTG	1377	Db	7604	TGCAAAACGAGGTTGGGAAGCAAGTCTCATCACTGCCAGCAGGACCCACTTCCAACCGA	7663
Db	6631	CAGGCCCCCTGCTGACACTCAGGAGTTTTCCTGCTCAGTGATGACCTCTTGTCTG	6690	QY	2458	CGGAGACCAAGTTTGTGATGGAAGAGCCATTTCTGAGTGGGACATCTTTGCTTCCAATG	2517
QY	1378	CTCATGCCACTGTGAAGGAAACAACAGTGTGAGTGAACCTGGATTAAGAGGTGACG	1437	Db	7664	CGGAGACCAAGTTTGTGATGGAAGAGCCATTTCTGAGTGGGACATCTTTGCTTCCAATG	7723
Db	6691	CTCATGCCACTGTGAAGGAAACAACAGTGTGAGTGAACCTGGATTAAGAGGTGACG	6750	QY	2518	GGATCATTTCTATGTCATTTTCAGGCTTTTAAAGCACCCCTGCCCGGTGACCTTGACC	2577
QY	1438	GAATCATGCAAGTTGTGGATTTCTGCAACAGGACAAACGGGGGCTGTGCAAGGTGG	1497				

[illegible]

RESULT 8

```

US-10-107-782-3
; Sequence 3, Application US/10107782
; Publication NO. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangoli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Miller, Isabelle,
; APPLICANT: Pothurajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1315)..(1315)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1316)..(1316)
; OTHER INFORMATION: Wherein n may be a, c, g or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1311)..(1313)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1314)..(1316)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-10-107-782-3
Query Match 91.9%; Score 2995.6; DB 16; Length 8495;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 3137; Conservative 0; Mismatches 10; Indels 116; Gaps

```

Qy	7	AGCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAAGTATTTCTTCCAGTTGCAGAGCAATT	66
Db	5311	ATCAGGAGCTTCCCAAGAACCCGAAAACTTCCCAAGTATTTCTTCCAGTTGCAGAGCAATT	5370
Qy	67	TCGTGAAGATCTGGTGGCCAGCCAGCCCTTCACTGTGTTTTTGGACCTTTATCTCAGCCT	126
Db	5371	TCGTGAAGATCTGGTGGCCAGCCAGCCCTTCACTGTGTTTTTGGACCTTTTATCTCAGCCT	5430
Qy	127	TTGATGAGGAAGCTCGGGTTAAAGACTGGGCAAAATACGGTTTAAATGCCCCAGGTTCTTC	186
Db	5431	TTGATGAGGAAGCTCGGGTTAAAGACTGGGCAAAATACGGTTTAAATGCCCCAGGTTCTTC	5490
Qy	187	GGTACCATTGGTTCGCCCTGCCACAGCTGTCTTGGAAAACTGAAATTTGATCTCAAAATG	246
Db	5491	GGTACCATTGGTTCGCCCTGCCACAGCTGTCTTGGAAAACTGAAATTTGATCTCAAAATG	5550
Qy	247	CTACTTCCCTCAAGGAGAGCCAAATAGTCACTCCGCTCTCTCAGAGCACGGTGATATAA	306
Db	5551	CTACTTCCCTCAAGGAGAGCCAAATAGTCACTCCGCTCTCTCAGAGCACGGTGATATAA	5610
Qy	307	ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATGTTTCATATCA	366
Db	5611	ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATGTTTCATATCA	5670
Qy	367	TAGACAAATTCCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGAA	426
Db	5671	TAGACAAATTCCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGAA	5730
Qy	427	TTCTGCAAAATCTTACGACTTTGGCAACAAACAAATGGCTACATAAAATTTAGCAACTTAA	486
Db	5731	TTCTGCAAAATCTTACGACTTTGGCAACAAACAAATGGCTACATAAAATTTAGCAACTTAA	5790
Qy	487	TACAGGACTCAGGTTTGGTGTGATGTATCAGGATCCCATCCACACCCAGTCACTCTCT	546
Db	5791	TACAGGACTCAGGTTTGGTGTGATGTATCAGGATCCCATCCACACCCAGTCACTCTCT	5850
Qy	547	TCGCGCCACCGACCAA-----GCCCTCCATGCCCTACCTGCTGAAACAAACAGGACT	597
Db	5851	TCGCGCCACCGACCAA-----GCCCTCCATGCCCTCCATGCCCTACCTGCTGAAACAAACAGGACT	5910
Qy	598	TCCTGTTCAACCAAGACAAAGGACCAAGCTGAAGGAGTATTTGAAAGTTTCATGTGATAC	657
Db	5911	TCCTGTTCAACCAAGACAAAGGACCAAGCTGAAGGAGTATTTGAAAGTTTCATGTGATAC	5970
Qy	658	GAGATGCCAAGGTTTAGCTGTGANTCTTCCCAATCCACTGCCCTGGAAGACCTGCAAG	717
Db	5971	GAGATGCCAAGGTTTAGCTGTGANTCTTCCCAATCCACTGCCCTGGAAGACCTGCAAG	6030
Qy	718	GTTCAGAGCTCAGTGTGAAATGTGGAGCTGGCAGGGAATCGGTGACCTTTTCTGAATG	777
Db	6031	GTTCAGAGCTCAGTGTGAAATGTGGAGCTGGCAGGGAATCGGTGACCTTTTCTGAATG	6090
Qy	778	GCCAAACTGCAGAAATTTGCAGCGGGAGCTCTTGTGACCTGGGTGGCCCTACGGCA	837
Db	6091	GCCAAACTGCAGAAATTTGCAGCGGGAGCTCTTGTGACCTGGGTGGCCCTACGGCA	6150
Qy	838	TTGACTGTCTGCTGATTCACACCTGGGGGGCGCTGTGACACCTTTTACTACTTTTCG	897
Db	6151	TTGACTGTCTGCTGATTCACACCTGGGGGGCGCTGTGACACCTTTTACTACTTTTCG	6210
Qy	898	ATGCTCGGGGAGTGTGGGAGCTGTGTCAATACTCCAGCTGCCCAAGGTGGAGTAAAC	957
Db	6211	ATGCTCGGGGAGTGTGGGAGCTGTGTCAATACTCCAGCTGCCCAAGGTGGAGTAAAC	6270
Qy	958	CAAGGGGTGACAGCAGAAAGTGTCTTACAACTGCCCTTCAAGAGGAACCTGGAAGGCT	1017
Db	6271	CAAGGGGTGACAGCAGAAAGTGTCTTACAACTGCCCTTCAAGAGGAACCTGGAAGGCT	6330
Qy	1018	GCCGGGAGCGGTGCAGCCTGGTGATPACAGATCCCCAGGTGTGCAAGGGCTACTTCGGGC	1077
Db	6331	GCCGGGAGCGGTGCAGCCTGGTGATPACAGATCCCCAGGTGTGCAAGGGCTACTTCGGGC	6390

QY	667	AGGTTTACGTGTGGATCTTCCCATCCACTGCCTCGAAGACCCCTGCAAGTTTCAGACG	726	QY	1747	AGTGCATGACAGACGCCAAATGTGTGACCTCCACTTCCAGGATACCACTGTGGGGTGT	1806
DB	2765	AGGTTTACGTGTGGATCTTCCCATCCACTGCCTCGAAGACCCCTGCAAGTTTCAGACG	2824	DB	3845	AGTGCATGACAGACGCCAAATGTGTGACCTCCACTTCCAGGATACCACTGTGGGGTGT	3904
QY	727	TGAGTGTGAAATGTGTGAGCTGGCAGGACATCGGTGACCTCTTTCTGAAATGGCCAAACCT	786	QY	1807	TCCATCTACGCTCCCACTGTGGGCGAGTATAAGCTGACCTTTTGACAAAGCCAGAGAGCCT	1866
DB	2825	TGAGTGTGAAATGTGTGAGCTGGCAGGACATCGGTGACCTCTTTCTGAAATGGCCAAACCT	2884	DB	3905	TCCATCTACGCTCCCACTGTGGGCGAGTATAAGCTGACCTTTTGACAAAGCCAGAGAGCCT	3964
QY	787	GCAGAAATGTGCAGCGGGAGCTCTTGTGTAACCTGGGTGTGGCTTACGGCATGATGTTC	846	QY	1867	GTGCCAAAGAAAGCTGGACCATGGCAACCTCAAAACAGCTCTCTCTATGCCAGAGGCCA	1926
DB	2885	GCAGAAATGTGCAGCGGGAGCTCTTGTGTAACCTGGGTGTGGCTTACGGCATGATGTTC	2944	DB	3965	GTGCCAAAGAAAGCTGGACCATGGCAACCTCAAAACAGCTCTCTCTATGCCAGAGGCCA	4024
QY	847	TGCTGATTCATCCCACTGTGGGGGCGCTGTGACACCTTTACTACTTTTCGATCGCTCGS	906	QY	1927	AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGTTGCTTACCCACACGCT	1986
DB	2945	TGCTGATTCATCCCACTGTGGGGGCGCTGTGACACCTTTACTACTTTTCGATCGCTCGS	3004	DB	4025	AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGTTGCTTACCCACACGCT	4084
QY	907	GGAGTGTGGGAGCTGTCTCAATACCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG	966	QY	1987	TGCTTCCAGAACTGTGTGGCTCTGTGTGGTGTGGGATAGTGGACTATGACCTAGACCCA	2046
DB	3005	GGAGTGTGGGAGCTGTCTCAATACCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG	3064	DB	4085	TGCTTCCAGAACTGTGTGGCTCTGTGTGGTGTGGGATAGTGGACTATGACCTAGACCCA	4144
QY	967	TGAAGCAGAGTGTCTCTACAACTGCTCCCTTCAAGAGGAACCTGGAAAGCTGCCGGAGC	1026	QY	2047	ACAAGAGTCAAAATGTGGGATGTCTTCTGTCTCGATGAAAGATGTGAATGCACCTGCA	2106
DB	3065	TGAAGCAGAGTGTCTCTACAACTGCTCCCTTCAAGAGGAACCTGGAAAGCTGCCGGAGC	3124	DB	4145	ACAAGAGTCAAAATGTGGGATGTCTTCTGTCTCGATGAAAGATGTGAATGCACCTGCA	4204
QY	1027	GGTGACGCTGGTGATACAGATCCCAAGTGTGCTGCAAGGGCTACTTCCGGGCGAGACTGTC	1086	QY	2107	AGGTGGGCTATGTGGGAGATGGCTTCTCATGTGAGTGGGAACTGTGTCAGGTCTCTGATGT	2166
DB	3125	GGTGACGCTGGTGATACAGATCCCAAGTGTGCTGCAAGGGCTACTTCCGGGCGAGACTGTC	3184	DB	4205	AGGTGGGCTATGTGGGAGATGGCTTCTCATGTGAGTGGGAACTGTGTCAGGTCTCTGATGT	4264
QY	1087	AGGCTGCCCTGGAGACACAGATGCCCGCTGTAATAACCGGGGTGTCTGCTTTGATCAGT	1146	QY	2167	CCTTCCCTCCTCACTCAAACTTCTGACCGGAAGTGTGGCTTCTCATGTGAGTGGGAACTGTGTCAGGTCTCTGATGT	2226
DB	3185	AGGCTGCCCTGGAGACACAGATGCCCGCTGTAATAACCGGGGTGTCTGCTTTGATCAGT	3244	DB	4265	CCTTCCCTCCTCACTCAAACTTCTGACCGGAAGTGTGGCTTCTCATGTGAGTGGGAACTGTGTCAGGTCTCTGATGT	4324
QY	1147	ACTCGGCGACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGGTGTGAGATGT	1206	QY	2227	GAGGCGTGCATTTCTTAGAACACCTGACTGACCTGCCATCCGCGGACCCCTCTTTGTGC	2286
DB	3245	ACTCGGCGACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGGTGTGAGATGT	3304	DB	4325	GAGGCGTGCATTTCTTAGAACACCTGACTGACCTGTCATCCGCGGACCCCTCTTTGTGC	4384
QY	1207	GCTGGCGGGGAGATTCGGGCTGATTCTGCTGCTGCTGCTGCTCCTCAGACACGGACAGT	1266	QY	2287	CACAGAACTAGTGGGCTGGGAGAAATGAGACCTTGTCTGGGGGGGACATCGAGCACACCC	2346
DB	3305	GCTGGCGGGGAGATTCGGGCTGATTGCTGCTGCTGCTGCTCCTCAGACACGGACAGT	3364	DB	4385	CACAGAACTAGTGGGCTGGGAGAAATGAGACCTTGTCTGGGGGGGACATCGAGCACACCC	4444
QY	1267	GCATGATGGCATCACGGGCTCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCT	1326	QY	2347	TGCGCAATGTAGCATGTTTTTCTACAAATGACCTTGTCTCAATGGCACCCCTTGGCAACGA	2406
DB	3365	GCATGATGGCATCACGGGCTCGGGCAGTGCCTCTGTGAAACGGGGTGGACAGGCCCT	3424	DB	4445	TGCGCAATGTAGCATGTTTTTCTACAAATGACCTTGTCTCAATGGCACCCCTTGGCAACGA	4504
QY	1327	CGTGTGACACTCAGGACGTTTTGCTGACGTGTGACGCTCTCTGTTCTGCTCATGCCA	1386	QY	2407	GGGTGGAAAGCAAGCTGCTCATCTGCCAGCCAGACCCACCTCCAAACCGACGAGACCA	2466
DB	3425	CGTGTGACACTCAGGACGTTTTGCTGACGTGTGACGCTCTCTGTTCTGCTCATGCCA	3484	DB	4505	GGCTGGAAAGCAAGCTGCTCATCTGCCAGCCAGACCCACCTCCAAACCGACGAGACCA	4564
QY	1387	CCTGTAAGAGAAACAACACGTGTGAGTGAACCTGGATATGAAGGTGACGGAAATCACAT	1446	QY	2467	GGTTTGTGATGGAAGAGCCATTTCTGCAGTGGGACATCTTTGCTTCCATGGGATCATTC	2526
DB	3485	CCTGTAAGAGAAACAACAAGTGTGAGTGAACCTGGATATGAAGGTGACGGAAATCACAT	3544	DB	4565	GGTTTGTGATGGAAGAGCCATTTCTGCAGTGGGACATCTTTGCTTCCATGGGATCATTC	4624
QY	1447	GCAAGTGTGGATTTCTGCAACACAGACAAACGGGGCTGTGCAAAAGGTGGCGAGATGCT	1506	QY	2527	ATGTCATTTCCAGGCTTTAAAGACACCCCTGCCCGGTGACCTTGACCCACACTGGCT	2586
DB	3545	GCAAGTGTGGATTTCTGCAACACAGACAAACGGGGCTGTGCAAAAGGTGGCGAGATGCT	3604	DB	4625	ATGTCATTTCCAGGCTTTAAAGACACCCCTGCCCGGTGACCTTGACCCACACTGGCT	4684
QY	1507	CCAGAGGGCACGAAGTCTCTGCGAGCTGCCAGAGGATACAAAGGGGACGGGACCA	1566	QY	2587	TGGGACGAGGATCTTTTGTCCATCATCTCTGTGACTGGGGCTGTTTGGCTTGGCTTGGCTT	2646
DB	3605	CCAGAGGGCACGAAGTCTCTGCGAGCTGCCAGAGGATACAAAGGGGACGGGACCA	3664	DB	4685	TGGGACGAGGATCTTTTGTCCATCATCTCTGTGACTGGGGCTGTTTGGCTTGGCTTGGCTT	4744
QY	1567	GCTGCAAGAGATAGACCCCTGTGCGACAGGCTTAAACGGAGGCTGTACGAGCAGGCCA	1626	QY	2647	ACTCTACTTTCCGATAAAACCGGAGAACCAATCGGCTTCCAGCATTTTGAATCGGAGAGG	2706
DB	3665	GCTGCAAGAGATAGACCCCTGTGCGACAGGCTTAAACGGAGGCTGTACGAGCAGGCCA	3724	DB	4745	ACTCTACTTTCCGATAAAACCGGAGAACCAATCGGCTTCCAGCATTTTGAATCGGAGAGG	4804
QY	1627	CCTGTAAGATGACAGGCGCGGGAGCAACAAGTGTGAGTGAATAAAGTCACTATGTGGAG	1686	QY	2707	ACATTAATGTTCAGCTCTTGGCAAGCAGCAGCTGAGATATCTCGAACCCCTTGTATG	2766
DB	3725	CCTGTAAGATGACAGGCGCGGGAGCAACAAGTGTGAGTGAATAAAGTCACTATGTGGAG	3784	DB	4805	ACATTAATGTTCAGCTCTTGGCAAGCAGCAGCTGAGATATCTCGAACCCCTTGTATG	4864
QY	1687	ATGGGCTGAATGTGAGCGGAGCAGCTGCCATTGACCGCTTACAGGACAAATGGGC	1746	QY	2767	AGAGCAAACTCAGCTCCCGGAGAACCTTTCTTACGACCCCTTTACGAGCTCTGAGAAC	2826
DB	3785	ATGGGCTGAATGTGAGCGGAGCAGCTGCCATTGACCGCTTACAGGACAAATGGGC	3844	DB	4865	AGAGCAAACTCAGCTCCCGGAGAACCTTTCTTACGACCCCTTTACGAGCTCTGAGAAC	4924
				QY	2827	GGCAGCTTGAGGGCAATGACCCCTTGGAGGACACTGTGA	2864


```
Db 4925 GGAGCTTGAGGCAATGACCCCTTGAGGACACTGTGA 4962
|||||
RESULT 10
US-09-842-930A-1
; Sequence 1, Application US/09842930A
; Publication No. US20020197681A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul
; TITLE OF INVENTION: Identification of Hyaluronan Receptor for Endocytosis
; FILE REFERENCE: 5820.603
; CURRENT APPLICATION NUMBER: US/09/842,930A
; CURRENT FILING DATE: 2001-04-22
; PRIOR APPLICATION NUMBER: 60/245,320
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/199,538
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4706
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-842-930A-1

Query Match 59.4%; Score 1936.8; DB 9; Length 4706;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 652; Indels 14; Gaps 6;

QY 11 GGAGCTTCCCAAGAACCCGAAACTTCCCGATTTCTTCCAGTTGCGAGGACATTTGCT 70
Db 1395 GGAGCTTCCCAAGAACCCCTTCCAGCTCCAGTACTTCTTCCAGTTGCGAGGACATTTGCT 1454
QY 71 GAAAGATCTGCTGGCGCCCGAGGCGCCCTTCACTCTTTTGGACCTTTATCTGCGACCTTTGA 130
Db 1455 CCGAGAGCTTGTGACCTTGGCCCTTCACTGCTTGGCGCTTTGTCTAGTCTCTTCA 1514
QY 131 TGAGGAAGCTCGGGTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGGTTCTTCGGTA 190
Db 1515 TCATGAGCCCGGATTAAGACTGGGATCAGCAGGCGCTCATGTCCAGGTTCTTCGCTA 1574
QY 191 CCATGCTGCTGCTGCGACCCAGCTGCTTCTGGAACAACTGAAATGATCTCAATGCTAC 250
Db 1575 TCACGTGCTGCTGCTGCGACCCAGCTGCTTGTGACAACTTAAAGTGAACCAAGTGCAC 1634
QY 251 TTCCCTCCAAAGAGAGCCAAATAGTATCATCTCCGTCTCTCAGACGCGGTATATAAATAA 310
Db 1635 GACCTTCCAAAGAGAGCCAGTTTCCATCTCTGCTCTCAGGACACTGTGTTCAATAAATA 1694
QY 311 TAAGGCTAAGATCATATCAAGTGATATCATCAGTACTAATGGATTTGTCATATCATAGA 370
Db 1695 TGAGGCAAGGTCCTGTCCAGTGACATCATCAGCACCAATGCGCTCATCCACGTTATAGA 1754
QY 371 CAAATGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGATTTCT 430
Db 1755 CAAGTTGCTGCTCCCAAAATTTGCTTATCACTCCCAAGATGCTTTGGCGAGGGTTCT 1814
QY 431 GCAAAATCTTACGACTTTGGCAACAAACAAATGGCTTACATCAAAATTTAGCAACTTAATA 490
Db 1815 GCAAAATCTTACTACAGTGGGAGCAACCAAGGATATACCAATTTACCAAGTTGATATA 1874
QY 491 GGACTCAGTTTGTGAGTGATCATCAGGATCCCATCCATCCACCCAGTCACTCTCTCTG 550
Db 1875 GGACTCAGGCTTGTGTCAGTCACTCACTGACTCCATCCATCCACCCAGTCACTGTTCTG 1934
QY 551 GCCACCCAGCAAGCCCTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 610
Db 1935 GCCTACGCAAGAACCCCTTGGAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1994
QY 611 AGACAACAGGACCAAGCTGGAAGGAGTATTTGAAGTTTCAATGATGATGATGATGATGAT 670
Db 1995 AGACAACAGGACCAAGCTGGAAGTCTTACTGAGTTCCAGTTCCAGTATCCAGGACTCC 2054

QY 671 TTTAGCTGTGGATCTTCCACATCCACTGCTGGAAGACCCCTGCAAGGTTTCAGAGCTGAG 730
Db 2055 TTTAGCTTCAGACCTCCCGAGGTTCTGCTTCTTGGAAAGACCCCTGCAAGGTTTCAGAGCTGAG 2114
QY 731 TGTGAAATGTGAGCTGTGGCAGGACATCGGTGACCTCTTTCTGAAATGGCAACCTGCGAG 790
Db 2115 TGTGAGGTGTGAAACTGGCAGTGACATCGGTGAGCTCTTTCTTAAACCAAAATGTGCGAG 2174
QY 791 AATTGTCAGCGGGAGCTCTTCTTGGACCTGGGTGGCTACGGATTCAGCTGTCTGCT 850
Db 2175 AATTCAACACCGGGGACTCTTCTTGGAGCTGGGTGGCTATGGCAATTCAGCTGTCTACT 2234
QY 851 GATTGATCCACCTTGGGGGCGGCTGTGACACCTTTACTTCTTCTGATGCTCGGGGA 910
Db 2235 CATGAATCTTACCTAGGTGGCGGATGTGACATTTTACTTCTTCTGATATTCGGGGGA 2294
QY 911 GTGTGGGAGCTGTGTCATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTGTGAA 970
Db 2295 GTGCGGAAAGTTGCAATTTTCACTCCCAAAATGCCCACTGAAGAGCAAGCCAAAGGGGTGA 2354
QY 971 GCAGAAGTGTCTCTACAACC---TGCCTTCAAGAGGAACTGGAGGCTGCGGGAGCG 1027
Db 2355 GAAGAAGTGTATCTACAACCCGCTTACCTTTCAGGAGGAACTGGAGGCTGCGAGACT 2414
QY 1028 GTGACGCTGTGTATACAGATCCCGAGGTGTGCAAGGCTACTTTCGGCGAGACTGTCA 1087
Db 2415 GTGACGCTGTGTATCCAAACCCCGAGGTGTGCGCATGTTTACTTTCATGCGAGACTGTCA 2474
QY 1088 GGCCTGCTGTGGAGACAGATGCCCCGTGTAAATAAACCGGGGTGTCTGCTTGTGATCAGTA 1147
Db 2475 GGCCTGCTGTGGAGACAGATACACCGGTGTAAACCAACCGGGGCACTGTGCGCGGATCTGTA 2534
QY 1148 CTCGCGCACCCGAGAGTGTAAATGCAACACCGGCTTCAATGGAGCGGCTGTGAGATGTG 1207
Db 2535 CACACCATGGACAGTGCCTATGTCACACCGGCTTCAACCGGGAAGGCTGCGAGCTCTG 2594
QY 1208 CTGGCGGGGAGATTCGGGCGCTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1267
Db 2595 CTGGCATGGAGATTTGGGCTGTGCTGTCAGCCCGCAGCTGCTCCGAGCATGGACAGTG 2654
QY 1268 CGATGATGGATACAGGCTCTCGGCGAGTGCCTCTGTGAAACGGGGTGGACAGGCGCTTC 1327
Db 2655 TGATGAGGGGATCACAGGCTCTCGGCGAGTGCCTCTGTGAAACAGGGTGGACAGCGCTTC 2714
QY 1328 GTGTGACACTCAGGAGTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1387
Db 2715 GTGTGACACTCCCAGCTGTATTGCGAGTGTGACACCTGCTGCTGCTGCTGCTGCTGCTG 2774
QY 1388 CTGTAAGGAGAACCAACAGCTGTGATGTAACTCTGGATTTATGAGTGTGACGGAATCACATG 1447
Db 2775 CTGTACGGAGAACCAACAGCTGTGTGTGTAACTTGTAACTTGTAACTTGTAACTTGTAACT 2834
QY 1448 CACAGTTGTGATTTCTGCAAAACAGGACCAACCGGGGCTGTGCAAAAGTGGCCAGATGCTC 1507
Db 2835 CACAGTCTGAGACTTCTGCAAAACAGGACCAACCGGGGCTGTGCGAAGGTGCTTAAAGTGTCT 2894
QY 1508 CCAGAAGGACAGGAGTCTCTCAGCTGCGGAGGAGATACAAAGGGGAGGCGGACAG 1567
Db 2895 CCAGAAGGACAGGAGTCTCTCAGCTGCGGAGGAGTCTCTCAGCTGCGGAGGAGTCTCAG 2954
QY 1568 CTGCACAGATAGACACCCCTGTGCGAGACGGGCTTAAACGGAGGGTGTCTACGAGACGCGAC 1627
Db 2955 CTGCATAGATAGACACCCCTGTGCGAGAGGTTGTCAACGGGGGATGCCATGAGCAGCGCAC 3014
QY 1628 CTGTAAGATGACAGGCGCGGCAAGACCAAGTGTGATGTGATGATGATGATGATGATGATG 1687
Db 3015 CTGCAAGATGACAGGCGCGGCAAGACCAAGTGTGATGTGATGTGATGTGATGTGATGTGAT 3074
QY 1688 TGGGCTCACTGTGACGCGGAGGAGCTGCCATTTGACGCTGCTTACAGGACATGGGCA 1747
Db 3075 CGGAGTGGACTGTGAGCTTACAGGAGCTGCGGCTGCGAGCTGCGGCTGCGAGTCCAGGAC 3134
```

QY 1748 GTGCCATCGACGACCAATATGTGCGACCTCCACCTTCAGGATACACATGTTGGGTGT 1807
DB 3135 GTGCCACCGAGATGCCAGCTGTGCAGACCTCTACTTTCAGGACACACCGTATGAGATAT 3194
QY 1808 CCATCTACCTCCCACTGGGCGAGTAAAGCTGACCTTTTGACAAAGCCAGAGGCGCTG 1867
DB 3195 CCATCTACCTCCCACTGGGCGAGTAAAGCTGACCTTTTGACAAAGCCAGAGGCGCTG 3254
QY 1868 TGCCAAAGAGCTGGGACCATGCGCAACCTACCAACAGCTCTCTATGCCCCAGAGGCCAA 1927
DB 3255 TGCCAAAGAGCTGGGACCATGCGCAACCTACCAACAGCTCTCTATGCCCCAGAGGCCAA 3314
QY 1928 GTACCACTGTCTCAGCAGGCTGGCTGAGACCGGCGGCTGCTACCCACAGCCTT 1987
DB 3315 GTATCACCTGTCTCGCGCGCTGGCTGGAGAGTGGCGGGTTCCTACCCGACTACGTA 3374
QY 1988 CGCCTCCCAAGAACTGTGGCTCTGGTGTGGTATGGGATAGTGAGCTATGGACCTTAGACCCAA 2047
DB 3375 TGCCCTCTCAGAAGTGTGGTGCAAACTGTTGGATCGTAGACTACGGATCCAGGGCCAA 3434
QY 2048 CAAAGATGAATGTGGGATGTCTCTGCTATCGGATGAAGATGTGAACCTGCACCTGCA 2107
DB 3435 CAAAGATGAATGTGGGATGTCTCTGTTACCGGATGAAGATGTGAACCTGCACCTGCA 3494
QY 2108 GTGGGCTATGTGGGATGGCTCTCTATGCACTGGGAACTGCTGAGTCTCTGATGTC 2167
DB 3495 GGAGGCTATGTGGGATGGCTCTCTGCTGAGTGGGAACTGCTGAGTCTCTGATGTC 3554
QY 2168 CTTCCCTCACTCACAACTTCTGACGGAAGTGTGGCTATTCACACAGCTCAGCTCG 2227
DB 3555 CTTCCCTCGCTCACAACTTCTGACGAGGTTGCTGGCTTTTCCAGAGCTCAGCCCG 3614
QY 2228 AGGCGTGATTTCTAGAACACTGACTGACCTGTCCATCCGCGGACACCTCTTTTGCC 2287
DB 3615 AGGACAGGCAATTTTGAACACCTGACTGACCTGTCCATCCGCTGGACACCTGTTTGCC 3674
QY 2288 ACAGAACAGTGGGCTGGGAGAAATGAGACCTTGTCTGGGCGGACATCGAGCACCACT 2347
DB 3675 ACAGAACAGTGGGCTACCGGGAATGAAGGCTGTCTGGCGGGAATTTGAGCACCACT 3734
QY 2348 CGCCAAATGTACGATGTTTTTCTACAAATGACCTTGTCAATGGCACCACTGCAACGAG 2407
DB 3735 CACTAATGTCAAGCTCTCTCTTTTACAAATGACCTTGTCAATGGTACCTTTCTGAGGACTAT 3794
QY 2408 GTGGGAAGCAAGCTCTCATCTCCAGCAGGACCCACTCCACCCGAGGAGCCAG 2467
DB 3795 GTGGGAAGCAAGCTCTCATCTACCTTCAGCCAGGACCACTCCACC--AAGACCCAG 3851
QY 2468 GTTGTGTGATGGAAGACCATCTGCACTGGGACATCTTGTGCTCCCAATGGGATCATTC 2527
DB 3852 GTTGTGTGATGGAAGATCATCTGCACTGGGACATCATCGCGCCCAATGGATCTCCCA 3911
QY 2528 TGTCAATTCAGGCTTTTAAAGCAACCCCTGCCCCGTGACCTTGACCCACACTGGCTT 2587
DB 3912 TATTATTTCTGAACCTTTGAGAGCTCTCTCCAGCGGACGAAACGGCTGCCCCACTCTGGCCT 3971
QY 2588 GGGAGAGGATCTCTTTTGCCATCATCTGTGTGACTGGGGCTGTGGCTTGTGCTGCTTA 2647
DB 3972 GGGAGAGGATATCTGTGCCCCGTGCTGCTGGTCACTGTGCGGATGCTCTGCGAGCTTA 4031
QY 2648 CTCCTACTTTCCGATAAACCGGAGAACAACTCCGCTTCAGCACTTTTGTAGTCGGAAGGA 2707
DB 4032 CTCCTACTTTCCGCTAAAGCAGCGAACCACTGCTTTTCCAGCGTTTTGA-TCAGAAGAGGA 4090
QY 2708 CATTAATGTGAGCTCTTGGGAGAGGAGCTGTGAGATATCTCGAACCCCTGTATGA 2767
DB 4091 CATTTGATGCTTGGCTTTTGGCAAGCAGAGCCCAAGATATCGCAACCCCTCTGTATGA 4150
QY 2768 GAGCACAACTCAGCTCCCGAGAACTCTCTACGACCCCTTCAACGAGCTCTGAAGAACG 2827
DB 4151 GACCTCAGCGCGGCAACCCGAGAGTCTCTCTGTGACCCCTTACAGACCCCTGGAGACA 4210
QY 2828 GCAGCTTGAGGGCAATGACCCCTTTGAGGACACTGTGAGGCGCTGGAGCGGAGATGCCAGC 2887

DB 4211 GGATCTGGAGGACGACGACCCCTCTGGGGCACTG--CGGTCTGACATGAGAAGCCAGC 4267
QY 2888 CATCACTACTCCCACTGGGCGCATCACTGTGAATTTCTCAGC--ACCAGTTGCCCTTT 2944
DB 4268 AAGCAACCACTGACGCTCCGCTCCAGCGTATCCAGCGCCCACTGTCTCATGATCAGTT 4327
QY 2945 AGGAACGTAAGTCTCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCTGGG 3004
DB 4328 GTTTTAAAGATGACAACTCATTAAGCCAGCCATACCTCACCTCTCTGTTAACTCGG 4387
QY 3005 GGTGCTTTCTGTGGGTGAGAGATGTTG-CTGTGCCCAACCCAGTACAGCTTCTCTCTCT 3063
DB 4388 ATTGTGCGCAGGCTAAGGAGCCATGTTGCTGTGATACCTGGGAGCTCCACCTCTCTCT 4447
QY 3064 GACCCCTTT 3071
DB 4448 GAGCCTAT 4455

RESULT 11
US-10-133-172-1
; Sequence 1, Application US/10133172
; Publication No. US20030104987A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: WEIGEL, JANET A
; TITLE OF INVENTION: METHODS OF USING THE HYALURONAN RECEPTOR FOR ENDOCYTOSIS
; FILE REFERENCE: 5864.014
; CURRENT APPLICATION NUMBER: US/10/133,172
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/286,468
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 09/842,930
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 4706
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-10-133-172-1

Query Match 59.4%; Score 1936.8; DB 15; Length 4706;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 652; Indels 14; Gaps 6;

QY 11 GGAGCTTCCAGAACCCGAAACTTCCAGATATTTCCAGTATTTCTCCAGTTCAGGAGCATTTCTGT 70
DB 1395 GGAGCTTCCAGAACCCCTTCGACGTCCTCCAGTACTTCTTCCAGTTCAGGAGCATGTGT 1454
QY 71 GAAAGATCTGTGCGGCCCGCCCTTCACTGTTTGTGACCTTTTATCTGACGCTTTGA 130
DB 1455 CGAGAGCTTGTGAGACCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCA 1514
QY 131 TGAGGAAGCTCGGGTTTAAAGACTGGGACAAATAACGGTTTAAATCCCGCAGGTTCTTCGGTA 190
DB 1515 TCATGAGCCCGGATTTAAAGACTGGGATCAGAGGGCCCTCATGTCCAGGTTCTTCGCTA 1574
QY 191 CATGTGCTGCTGCGCCAGCAGCTGTTCTGGAAACCTGAATTCATCTCAATGCTAC 250
DB 1575 TCAGTGTGGGTGCGCAGCAGCTGTTGTGGCAACCTTAAAGTACCAAGTGCAC 1634
QY 251 TTCCCTCCAGGAGGACCAATAGTCTATCTCCGCTCTCTCAGAGCACCGGTGTATATAATAA 310
DB 1635 GACCTCCAGAGGAGCCAGTTTCCATCTCTGCTCTCAGGACACTGTGTTCATATAACAA 1694
QY 311 TAAGGTATAGATCATATCCAGTGAATCATCATGATCATATATGGGATTTGTTATATCATAGA 370
DB 1695 TGAGGCGAAGGTCTCTGCTCCAGTGCATCATCAGCACCAATGGGCTCATCCACGTTATAGA 1754
QY 371 CAAATGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACCTCTGGAAGATTTCT 430

Db 1755 CAAGTTGCTGTCTCCCAAAACCTTGCTTATCACCCCAAGATGCTTTGGGACAGGGTTCT 1814
QY 431 GCAAAATCTTACGACTTTGGCAACAAATGGCTACATCAAAATTTAGCAATTAATACA 490
Db 1815 GCAAAATCTTACTACAGTGGCAGCAAAACCAAGATATACCAATTCAGCAAGTTGATACA 1874
QY 491 GCACTCAGGTTTGTGAGTGTGATACCGATCCATCCATCCACACCCAGTCACTCTCTTG 550
Db 1875 GCACTCAGGTTTGTGAGTGTGATACCGATCCATCCACACCCAGTCACTCTCTTG 1934
QY 551 GCCCAGCAGCAGCCCTCCATGCCCTACTCTGAACAAAGGACTTCTCTTCAACA 610
Db 1935 GCCTAGGACAAAGCCCTGGAGCCCTTGCCCAAGACAGCAGGAGCTTCTCTTCAATCA 1994
QY 611 AGACAACAGGACAAAGCTGAAGGAGTATTGAAAGTTTTCATGTGATACAGAGTCCCAAGT 670
Db 1995 AGACAACAGGACAAAGCTGAAGTCTTACCTGAAGTTCCACGTGATCCGAGACTCCCAAGC 2054
QY 671 TTTAGCTGTGGATCTTCCACATCCACTGCCCTGGAGACCCCTGCAAGGTTCAAGGTGAG 730
Db 2055 TTTAGCTTCAGACCTCCCAAGTCTGCTTCTGGAAGACCCCTGCAAGGCTCAGAGCTGAG 2114
QY 731 TGTGAAATGTGAGCTGGCAGGACATCGGTGACCTCTTTCTGAAATGSCCAACCTGCGAG 790
Db 2115 TGTGAGGTGTGAACTGGCAGTGACATGGTGAGCTCTTTTAAACGAACAAATGTGCAG 2174
QY 791 AATTGTGAGCGGGAGCTCTTGTGTTGACCTGGGTGTGGCTTACGGCATTTGACTGTCTGT 850
Db 2175 ATTATACACCGGGAGCTCTTGTGAGCTGGGTGTGGCTTATGSCATTTGACTGTCTACT 2234
QY 851 GATTGATCCCACTGGGGGCGCTGTGACACTTTACTATTCTGATTCGCTCGGGGGA 910
Db 2235 CATGAATCTTACCTAGGTGGCGGATGTGACACTTTTACTACTCTCGATATTCCGGGGA 2294
QY 911 GTGCGGAGCTGTGTCAATCTCCAGCTGCCAAGGTGGAGTAAACCAAGGGGTGCA 970
Db 2295 GTGCGAAGTTGCATTTTCACTCCCAATGGCCACTGAAGCAAGCCAAAGGGGTGAA 2354
QY 971 GCAGAAGTGTCTCTAAACCC---TGCCCTTCAAGAGGAACCTTGGAGGCTGCGGGGCG 1027
Db 2355 GAAGAAGTGTATCTAAACCCGTTACCTTTTCAAGAGGAACCTTGGAGGCTGCCAGAACCT 2414
QY 1028 GTGAGCTGTGTGATACAGATCCCAAGTGTGCAAGGCTACTTTCGGGGGAGACTGTCA 1087
Db 2415 GTGACACGTGTGTATCCAAACCCCAAGTGTGCAAGGCTACTTTCATGCCAGACTGTCA 2474
QY 1088 GGCCTGCCCTGGAGGACAGATGCCCGGTGTAATAACCGGGGTGTCTGCCTTGTATCACTA 1147
Db 2475 GGCCTGCCCTGGAGGACAGATACACCGTGTAAACCGGGGCACTGTGCCGGATCTGTA 2534
QY 1148 CTGGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGAGATGT 1207
Db 2535 CACACCAATGGACAGTGCCTATGCCACACCGGCTTCAACGGGACAGCCTCGAGCTCTG 2594
QY 1208 CTGGCGGGGAGATTCGGGCTGTATGTCTGCCCTGTGGCTGTCTGACACAGGACAGTG 1267
Db 2595 CTGGCATGGGAGATTTGGGCTGTGCTGTGACGCCCGGAGCTGTCTCGAGCATGGACAGTG 2654
QY 1268 CGATGATGGCATCACGGGCTCGGGGCTGCTCTGTGAACCGGGGTGGACAGGCCCTC 1327
Db 2655 TGATGAGGGGATCACAGGCTCGGGGAGTGCTCTGTGAAACAGGGTGGACAGCGGCTTC 2714
QY 1328 GTGTGACACTCAGGAGTTTTGCCCTGCGAGTGTGACGCTCTCTTGTCTGCTATGCCAC 1387
Db 2715 GTGTGACACTCCACAGCTGTATTTCGAGTGTGCACACCTGTTGCTCCGTCAGCCAC 2774
QY 1388 CTGTAAAGAGACACAGTGTGAGTGTAACTCGATTATGAAGTGTGCGGAATCAATG 1447
Db 2775 CTGTACGGAGAACACAGCTGTGTGTAACTTGAACCTACGAAGGTGACGGGATCAATG 2834
QY 1448 CACAGTGTGTGATTTCTCAAAACAGCAACCGGGGCTGTCAAGGTGGCCAGATGCTC 1507
Db 2835 CACAGTGTGTGACTTCTGCAAAACAGCAACCGGGGCTGTGCGAAGGTGCTTAAGTGTCT 2894

QY 1508 CCAGAAGGGCACAAAGGTCTCTCTGAGCTGCCAGAAGGATACAAAGGGGACGGGCACAG 1567
Db 2895 CCAGAAGGGCACCCAGTCTCTTTCAGCTGCAAGAAAGGCTACAAAGGGGATGGCTACAG 2954
QY 1568 CTGCAACAGATAGACCCCTGTGACAGCGGCTTTAACGGAGGGTGTCAAGCAGCCAC 1627
Db 2955 CTGCATAGATAGACCCCTGTGACAGCGGCTTCAACGGGGGATGCCATGACACGCCAC 3014
QY 1628 CTGTAAGATGACAGCGCGGCAAGCAACAGTGTGAGTGTAAAGTCACTATGTGGAGA 1687
Db 3015 CTGCAGSAGTACGGGCGCCAGGCAAGCATAGTGTGAATGTAAAGTCACTATGTGGGGA 3074
QY 1688 TGGGCTGAACTGTGAGCGGAGCAGCTGCCAATTCACCGCTCTTACAGGACAAATGGGCA 1747
Db 3075 CGGAGTGGACTGTGAGCTGTGAGCAGCTGCCCTCGACCGTGTCTTACAGGACAAAGGCA 3134
QY 1748 GTGCCATGACAGCGCAAAATGTGTGACCTTCCACTTCCAGGATACCACTTTGGGGTGT 1807
Db 3135 GTGCCACCCAGATGCGAGCTGTGCAGACCTCTACTTCCAGGACACGACCGTAGGAGTAT 3194
QY 1808 CCATCTACGCTCCCACTGGGCGCAGTATAAGCTGACCTTGCACAAAGCAGAGAGCCCTG 1867
Db 3195 CCATCTACGCTCCCACTGGGCGCAGTACAACTGACATTTGACAAAGCCAAAGAGCCCTG 3254
QY 1868 TGCCAAACGAAGCTGCGACCATGGCAACCTACAACAGCTCTCTATGCCAGAGGCCAA 1927
Db 3255 TGCCAAAGAGCTGGACCATAGCCACCTACAACAGCTCTCTATGCCAGAGGCCAA 3314
QY 1928 GTACACCTGTGTGACAGCTGGCTGGAGACCGGGGGTGTCTACCCACAGCCCT 1987
Db 3315 GTATCACCTGTGTGCGCGCTGGCTGGAGAGTGGCGGGTGTCTACCCGACTACGTA 3374
QY 1988 CGCCTCCAGAACCTGTGGCTGTGGTGTGGTGGAGTAGTGACATGACACTAGACCCAA 2047
Db 3375 TGCTCTCAGAAAGTGTGGTGCAACGTTGTTGGGATCTGACTACGATCCAGGCCAA 3434
QY 2048 CAAGAGTGAATGTGGAGTGTCTTCTGCTATCGAATGAAGATGTGAATGCACCTGCA 2107
Db 3435 CAAGAGTGAATGTGGAGTGTCTTCTGTTACGGATGAAGATGTGAATGCACCTGCA 3494
QY 2108 GGTGGCTATGTGGAGATGGCTTCTCATGAGTGGGAACTGTGTGAGGTCTGTAGTGC 2167
Db 3495 GGCAGCTATGTGGAGATGGCTTCTGTGAGTGGGAACTGTGTGAGGTCTCATGTC 3554
QY 2168 CTTCCCTCACTCAAACTTCTGACGGAAGTGTGCTTATTCACACAGCTCAGCTG 2227
Db 3555 CTTCCCTCGCTCAAACTTCTGACAGAGGTGTGCTTTTCCAAAGGCTCAGCCG 3614
QY 2228 AGGCCGTGCAATTTCTAGAACACCTGACTGACTGTCCATCGCGGACCCCTCTTTGTGCC 2287
Db 3615 AGGACAGCAATTTTGAACACACCTGACTGACTGTCCATCGTGGCACCCTGTTGTGCC 3674
QY 2288 ACAGAACAGTGGGCTGGGGAGAAATGAGACTGTGTGGGGGGGACATTCGAGACCACT 2347
Db 3675 ACAGAACAGTGGGCTACCGGGAATAAGAGCTGTGTGGCGGGGACATTCGAGACCACT 3734
QY 2348 CGCCAAATGTGACAGATGTTTTCTCAATGACTGTGTAAGTGGACCCCTGCAACAG 2407
Db 3735 CACTAATGTCAACGTCTCTTTTAAATGACTGTGTAAGTGGTACCTTTCTGAGGACTAT 3794
QY 2408 GGTGGGAGCAAGCTGTCTCATCTGCGCAGCCAGGACCCACTCCAAACGACGAGACCA 2467
Db 3795 GCTGGGAGCAACTGTCTCATTTACCTTCAGCCAGGACGAGCTCCACC---AAGAGACCA 3851
QY 2468 GTTTGTTGATGGAAAGCCATTTGACAGTGGGACATCTTTGCTCCCAATGGGATCATCA 2527
Db 3852 GTTTGTTGATGGAAAGTCCATTTGACAGTGGGACATCATCGCCGCAATGGAATCCCTCA 3911
QY 2528 TGTCAATTTCCAGGCTTTAAAGCACCCCTGCCCCCGTGTGACCTTTGACCCACACTGGCT 2587
Db 3912 TATTATTCTGAACCTTTGAGAGCTCTCTCCCAAGGACAGGCTCGGCTCCCACTCTGGCT 3971

1596	Db	CAAGGCTGGGCAAGGGCTGCCACCGCAATTGTGTACACCACTCTGGGAAGCCAGCTGC	1655
1059	QY	TGCAAGGGCTACTTTCGGGGCAGAGCTGTCAAGGCTGCGCTGGAGAGACAGATGCCCGTCT	1118
1656	Db	TGCGCTGGTCACTATGCGAGTGAGTGCACAGCTTGGCCCTGGCGGCCACAGACCCCTTGT	1715
1119	QY	AATAACCGGGGTGTCTGCTTGATFCAGTACTCGGCCACCGAGAGGTAAATGCAACAC	1178
1716	Db	AGTGACGTGGCTGTGCAATGACCGGCATGAGTGGCAGTGGCGAGTGTCTGTGCCGTCA	1775
1179	QY	GGCTTCAATGGGACCGGCGTGTGAGATGTGTGGCCGGGGAGATTCGGGCGCTGATGTCTG	1238
1776	Db	GGTTTTGTGGGACAGCGCTGTGAACCTGTGTCTCTGTGCTCTGTGCTCTTGGGCCCCATTGTCAA	1835
1239	QY	CCCTGTGGCTGCTCAGACCAAGGACAGTGCAGATGGCATCAGGGCTCCGGGCAAGTGC	1298
1836	Db	GCCTGCCGCTGCATGTGTGCATTGGCGCGTGTGATGAGGGCCCTGGGGGCTCTGGCTCCTGC	1895
1299	QY	CTCTGTGAAACGGGTGGACAGGCGCCCTGTGTGCACACTCAGGAGAGTTTGGCTCGAGTG	1358
1896	Db	TTCTGTGATGAAGGCTGACCTGGGCGACGCTGTGAGGTGCAACTGAGAGTGCAGCCTGTG	1955
1359	QY	TGTAGCCCTCCTTGTTCGTCTCATGCCACCTGTAAAGGAGAACACGTTGTGAGTGTAAAC	1418
1956	Db	TGTACCCCAACCTGTGCACCCGAGGCTGTGTGCCGTGCAGGCAACAGCTGTGAGTGCAGC	2015
1419	QY	CTGATTTATGAAGGTGACGGAAATCACATGCACAGTTGTGGATTTCTTCGAAACAGGACAAC	1478
2016	Db	CTGGGCTATGAAGGGGATGGGCGTGTGTGTACAGTGGCAGACCTGTGCCAGGACGGGCAT	2075
1479	QY	GGGGCTGTGCAAGGTGGCGAGATGCTCCAGAAAGGGCACGAAGTCTCCTCGAGCTGC	1538
2076	Db	GGTGGCTGCAGTAGACAGCCCAACTGTAGCCAGGTAGGAACAATGCTCACTTTGTAACCTGC	2135
1539	QY	CAGAAGGGGATACAAAGGGGACGGGACAGCTTGCAACAGATAGACCCCTGTGCAGACGGC	1598
2136	Db	CTGCCCGACTACGAGGGTGTATGGCTGGAGCTGCCGGCCGCCAACCCCTGCACAGATGGC	2195
1599	QY	CTTAAACGAGGGTGTACAGACAGCCACCTGTGAAGATGACAGGCCCGGGCAGGACACAAG	1658
2196	Db	CACCGCGGGGCTGCAGCGAGCAGCCCAACTGCTTGAGCACCGGCGCTGAAACACACGGCGC	2255
1659	QY	TGTGAGTGTAAAGTCACTATGTGCGAGATGGGTGAACCTGT --- GAGCCGGAGCAGCTG	1715
2256	Db	TGTGAGTGCACGAGGCTTACGTAGCGATGCACTGCAGTGTCTGGAGGAGTTCGGAACCA	2315
1716	QY	CCCATGTACCGCTGCTTACAGGACAAATGGGAGTGCATGCAGACGCGCAATGTGTGCAC	1775
2316	Db	CTGTGTGACCGCTGCTTGGGCCAGCACCGCCCTGCCACTCAGATGCCATGTGTGACTGAC	2375
1776	QY	CTTCACTTCCAGGATACCACTGTGGGGTGTTCATCTACGCTTCCCACTGGGGCCAGTAT	1835
2376	Db	CTTGCACTTCCAGGAGAAACGGGCTGGCGCTTTTCCACCTCCAGGCGCACGCGCGCCCTTAT	2435
1836	QY	AAGCTGACCTTTGACAAAGCCAGAGAGGCGCTGTGCCAAGCTGCGACCATGCGCAACC	1895
2436	Db	GGTCTGAACCTTTCGAGGCTGAGCGGCGATTCGAGCAACAGGAGCGCTCTTGCTTCA	2495
1896	QY	TACAAACCAAGCTCTCCTATGCCAGAGGCCAAGTACCACTGTGCTCAGCAGGCTGGCTG	1955
2496	Db	TTCCCTCAGCTCTCTGCTGCCAGCAGCTGGGCTTCCACCTGTGCTCATGGGCTGGCTG	2555
1956	QY	GAGACGGGCGGGTTCCTTACCCCAAGCCTTGGCTCCCAAGAACTGTGGCTCTGGTGTG	2015
2556	Db	GCCAATGGCTCCACTGCCCAACCCCTGTGGTTTTTCCCTGTGGCGGAGCTGTGGCAATGTGCGG	2615
2016	QY	GTTCGGATAGTGGACTATGGACCTAGACCCCAACAGAGTGAATGTGGGATGCTTCTGCG	2075
2616	Db	GTGGGCATAGTCAGCCCTGGGTCGCCGAGAACCTCTTCAGAAACGCTGGGATGCTACTGCG	2675
2076	QY	TATCGGATGAAAGATGTGAACCTGCACCTGCAAGGTGGGCTATGTGGGAGATGG --- CTTTC	2132

```

RESULT 15
US-09-918-995-30654
; Sequence 30654, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30654
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(448)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30654

```

	Query Match	12.5%	Score 408.8	DB 10	Length 448
	Best Local Similarity	98.1%	Pred. No. 4.2e-118		
	Matches 413	Conservative 0	Mismatches 8	Indels 0	Gaps 0
QY	1186	ATGGGACGGCGTGTGACATGTCGTGGCCGGGGAGATTGGGCCCTGATTGTCGCCCTGTG	1245		
DB	28	ATCGGACGGCGCGNAGATGTGTCGTGGCCGGGGAGATTGGGCCCTGATTGTCGCCCTGTG	87		
QY	1246	GCTGCTCAGACACGGACAGTGCATGATGSCATCACGGGCTCCGGGCAGTGCCCTCTGTG	1305		
DB	88	GTTGCTCAAAACACGGACTGTGCCATGATGGCATCACGGGCTCCGGGCAGTGCCCTCTGTG	147		
OY	1306	AAACGGGTGGACAGGCCCTCTGTGTGACACTCAGGCAGTTTTTCCTCCAGTGTGTACGC	1365		

Db	148	AAACGGGGTGACAGGCCCTCGTGTGACACTCAGGCAGTTTGTGCTGCAGTGTACGC	207
Qy	1366	CTCCTTGTCTCTCATGCCACTGTAAAGGAGAACACACGTCGTGAGTGTAACTTGGATT	1425
Db	208	CTCCTTGTCTCTCATGCCACTGTAAAGGAGAACACACGTCGTGAGTGTAACTTGGATT	267
Qy	1426	ATGAGGTGACGGGATCATGTCACAGTTGTGGATTCTGCAACAGGACAAACGGGGGCT	1485
Db	268	ATGAGGTGACGGGATCATGTCACAGTTGTGGATTCTGCAACAGGACAAACGGGGGCT	327
Qy	1486	GTGCAAGGTGGCCAGATGCTCCCAAGAGGCAACGAGTCTCCTGCAGCTGCCAGAGG	1545
Db	328	GTGCAAGGTGGCCAGATGCTCCCAAGAGGCAACGAGTCTCCTGCAGCTGCCAGAGG	387
Qy	1546	GATACAAAGGGGACGGGCACAGCTGCACAGATAGACCCCTGTGTCAGACGGCCTTAACG	1605
Db	388	GATACAAAGGGGACGGGCACAGCTGCACAGATAGACCCCTGTGTCAGACGGCCTTAACG	447
Qy	1606	G 1606	
Db	448	G 448	

Search completed: June 21, 2004, 16:49:26
Job time : 1332 secs

Pred. No. is the number of results predicted by chance to have a

RESULT 1
AX686401
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCES
AUTHORS

AX686401	LOCUS	AX686401	3260 bp	DNA	linear	PAT 29-MAR-2003
	DEFINITION	Sequence 210 from Patent WO02059315.				

KEYWORDS
SOURCE
Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Malyankar, U., Shenoy, S., Spytek, K.A., Gangolli, E., Miller, C., Boldog, F., Li, L., Taupier, R.J., Kekuda, R., Smithson, G.,

Zerhusen, B. D., Liu, X., Colman, S. D., Tchernev, V., Shi, J., Edinger, S.,
Stone, D., Sciore, P., Millet, I., and Rothenberg, M.
Human nucleic acids and polypeptides and methods of use thereof
Patent: WO 02059315-A 210 01-AUG-2002;
Curagen Corporation (US)

FEATURES
Location/Qualifiers
1..3260
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN	Query Match 100.0%; Score 3259; DB 6; Length 3260; Best Local Similarity 100.0%; Pred. No. 0; Matches 3260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	GGCAGCAGCAGGAGCTTCCCAAGAACCCGAAACCTTCCAGTATTTCTTCCAGTTGCAGG	60			
Db	1	GGCAGCAGCAGGAGCTTCCCAAGAACCCGAAACCTTCCAGTATTTCTTCCAGTTGCAGG	60			
QY	61	AGCATTTCTGTAAGATCTGGTCGGCCAGGCCCTTCACTGTTTTCACCTTTATCTG	120			
Db	61	AGCATTTCTGTAAGATCTGGTCGGCCAGGCCCTTCACTGTTTTCACCTTTATCTG	120			
QY	121	CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCAGG	180			
Db	121	CAGCCTTTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCAGG	180			
QY	181	TTCTTCGGTACCATTGTGCTCGCTGCCACAGCTGCTTCTGGAACCTTGAATTTGATCT	240			
Db	181	TTCTTCGGTACCATTGTGCTCGCTGCCACAGCTGCTTCTGGAACCTTGAATTTGATCT	240			
QY	241	CAAAATGCTACTTCCCTCCAGGAGAGCCAAATAGTCACTCCGCTCTCAGAGCAAGGTTG	300			
Db	241	CAAAATGCTACTTCCCTCCAGGAGAGCCAAATAGTCACTCCGCTCTCAGAGCAAGGTTG	300			
QY	301	ATATAAATAATAGGCTAAGATCATATCCAGTGTATCATCAGTACTCAATCGGATTTGTC	360			
Db	301	ATATAAATAATAGGCTAAGATCATATCCAGTGTATCATCAGTACTCAATCGGATTTGTC	360			
QY	361	ATATCATAGACAAATTTGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTG	420			
Db	361	ATATCATAGACAAATTTGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTG	420			
QY	421	GAAGAAATCTGCAAAATCTTACGACTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCA	480			
Db	421	GAAGAAATCTGCAAAATCTTACGACTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCA	480			
QY	481	ACTTAATACAGGACTCAGGTTTGTGAGTGTATCATCCGATCCCATCCACACCCCAAGTCA	540			
Db	481	ACTTAATACAGGACTCAGGTTTGTGAGTGTATCATCCGATCCCATCCACACCCCAAGTCA	540			
QY	541	CTCTCTTCTGGCCCAACGACCAAGCCCTCCATGCTCCCTTACCTGCTGAACAAACAGGACTTCC	600			
Db	541	CTCTCTTCTGGCCCAACGACCAAGCCCTCCATGCTCCCTTACCTGCTGAACAAACAGGACTTCC	600			
QY	601	TGTTCAACCAAGACAAAGGACCAAGCTGAAGAGTATTTGAAGTTTCATGTGATACGAG	660			
Db	601	TGTTCAACCAAGACAAAGGACCAAGCTGAAGAGTATTTGAAGTTTCATGTGATACGAG	660			
QY	661	ATGCCAAGGTTTATAGTGTGGATCTTCCCAATCCATCCATCCATCCATCCATCCATCCAT	720			
Db	661	ATGCCAAGGTTTATAGTGTGGATCTTCCCAATCCATCCATCCATCCATCCATCCATCCAT	720			
QY	721	CAGAGCTGAGTGTGAATGTGGAGCTGGGAGGACATCGGTGACCTCTTTCTGAATGGCC	780			
Db	721	CAGAGCTGAGTGTGAATGTGGAGCTGGGAGGACATCGGTGACCTCTTTCTGAATGGCC	780			
QY	781	AAACCTGCAGAAATTTGTGACGGGAGCTCTTTGTTGACCTGGGTGGCCCTACGGCATTTG	840			
Db	781	AAACCTGCAGAAATTTGTGACGGGAGCTCTTTGTTGACCTGGGTGGCCCTACGGCATTTG	840			
QY	841	ACTGTCTGCTGATTGATCCCAACCTTGGGGGGCCGCTGTGACACCTTTACTCTTTCGATG	900			

DB	841	ACTGTCTGCTGATTGATCCACCTTGGGGGGCCGCTGTGACACCTTTACTACTTTCGATG	900					
QY	901	CCTCGGGGAGTGTGGGAGCTGTGTCAATATCTCCAGAGTCCCAAGGTGGAGTAAACCA	960					
DB	901	CCTCGGGGAGTGTGGGAGCTGTGTCAATATCTCCAGAGTCCCAAGGTGGAGTAAACCA	960					
QY	961	AGGGTGTGAAGCAGAGTGTCTCTAACCTTCCCTTCAAGAGGAACTTGAAGGCTGCC	1020					
DB	961	AGGGTGTGAAGCAGAGTGTCTCTAACCTTCCCTTCAAGAGGAACTTGAAGGCTGCC	1020					
QY	1021	GGGAGCGGTGACAGCCTGTGTATACAGATCCCCAGGTGTCTCAAGGCTACTTTCGGCGAG	1080					
DB	1021	GGGAGCGGTGACAGCCTGTGTATACAGATCCCCAGGTGTCTCAAGGCTACTTTCGGCGAG	1080					
QY	1081	ACTGTACAGCCTCCCTGGAGGACAGATGCCCGCTGTAAATAACCGGGGTGTCTGCCTTG	1140					
DB	1081	ACTGTACAGCCTCCCTGGAGGACAGATGCCCGCTGTAAATAACCGGGGTGTCTGCCTTG	1140					
QY	1141	ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG	1200					
DB	1141	ATCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTG	1200					
QY	1201	AGATGTGCTGGCCGGGAGATTTCGGGCTGTGTTGTCTGCCCTGTGCTCTCAGACCAAG	1260					
DB	1201	AGATGTGCTGGCCGGGAGATTTCGGGCTGTGTTGTCTGCCCTGTGCTCTCAGACCAAG	1260					
QY	1261	GACAGTGGATGATGGCATCACGGGCTCGGGAGTGTCTGTGAAACGGGGTGGACAG	1320					
DB	1261	GACAGTGGATGATGGCATCACGGGCTCGGGAGTGTCTGTGAAACGGGGTGGACAG	1320					
QY	1321	GGCCCTCGTGTGACACTCAGGAGTGTCTGCTGAGTGTGACGCTCTCTTCTGTCTC	1380					
DB	1321	GGCCCTCGTGTGACACTCAGGAGTGTCTGCTGAGTGTGACGCTCTCTTCTGTCTC	1380					
QY	1381	ATGCCACCTGTAAAGAGAAACAACGCTGTGAGTGTAACTGGATTATGAAGGTGACGAA	1440					
DB	1381	ATGCCACCTGTAAAGAGAAACAACGCTGTGAGTGTAACTGGATTATGAAGGTGACGAA	1440					
QY	1441	TCACATGACAGTGTGGATTTCTGCAACAGGACAAACGGGGCTGTGCAAGGTGGCCA	1500					
DB	1441	TCACATGACAGTGTGGATTTCTGCAACAGGACAAACGGGGCTGTGCAAGGTGGCCA	1500					
QY	1501	GATGTCTCCAGAGGGACGAGGCTCTCCTGAGCTGTCCAGAGGATACAAAGGGGACG	1560					
DB	1501	GATGTCTCCAGAGGGACGAGGCTCTCCTGAGCTGTCCAGAGGATACAAAGGGGACG	1560					
QY	1561	GGCAGAGTGTGACAGATAGACCCCTGTGCTGAGACGGCCCTTAAACGGAGGCTGTACAGC	1620					
DB	1561	GGCAGAGTGTGACAGATAGACCCCTGTGCTGAGACGGCCCTTAAACGGAGGCTGTACAGC	1620					
QY	1621	ACGCCACCTGTAAAGATGACAGGCGCGGCAAGCACAAGGTGAGTGTAAAGTCACTATG	1680					
DB	1621	ACGCCACCTGTAAAGATGACAGGCGCGGCAAGCACAAGGTGAGTGTAAAGTCACTATG	1680					
QY	1681	TGGAGATGGGCTGAACTGTGAGCCGGAGCAGTGTCCCATTTGACCCCTGTCTACAGACA	1740					
DB	1681	TGGAGATGGGCTGAACTGTGAGCCGGAGCAGTGTCCCATTTGACCCCTGTCTACAGACA	1740					
QY	1741	ATGGGAGTGTGCAATGACAGCGCCAAATGTGTGCACTTCCAGGATACACCTGTG	1800					
DB	1741	ATGGGAGTGTGCAATGACAGCGCCAAATGTGTGCACTTCCAGGATACACCTGTG	1800					
QY	1801	GGGTGTTCCATCTTACGCTCCCACTGGCCAGTATTAAGTGTACCTTTGCAAGCCAGAG	1860					
DB	1801	GGGTGTTCCATCTTACGCTCCCACTGGCCAGTATTAAGTGTACCTTTGCAAGCCAGAG	1860					
QY	1861	AGGCTGTGCGCAACGAACTGTGCGCAACCTTACAAACAGCTCTCTATGCGCAGA	1920					
DB	1861	AGGCTGTGCGCAACGAACTGTGCGCAACCTTACAAACAGCTCTCTATGCGCAGA	1920					
QY	1921	AGGCCAAGTACCACTGTCTCAGCAGGCTGGTGGAGACCGGGGGTGTCTTACCCCA	1980					

Db 1921 AGGCCAAGTACCACTGTGCTCAGCAGCGCTGGCTGGAGACCGGGCGGGTTCCTACCCCA 1980
Qy 1981 CAGCCTTCGCTCCCGAAGATGTGGCTCTGGTGTGGGATGAGTGGACCTA 2040
Db 1981 CAGCCTTCGCTCCCGAAGATGTGGCTCTGGTGTGGGATGAGTGGACCTA 2040
Qy 2041 GACCAACAAGAGTGAAGATGTGGATGTCTCTGCTATCGGATGAAGATGAAGTGA 2100
Db 2041 GACCAACAAGAGTGAAGATGTGGATGTCTCTGCTATCGGATGAAGATGAAGTGA 2100
Qy 2101 CCGCAAGTGGGCTATGTGGAGATGGCTTCTCATGCAAGTGGGAACCTGCTGCAGTCC 2160
Db 2101 CCGCAAGTGGGCTATGTGGAGATGGCTTCTCATGCAAGTGGGAACCTGCTGCAGTCC 2160
Qy 2161 TGATGCTTCCCTCCTCACTCAAACTTCTGCTGCAAGTGGGCTGCTGCTGCTGCTGCT 2220
Db 2161 TGATGCTTCCCTCCTCACTCAAACTTCTGCTGCAAGTGGGCTGCTGCTGCTGCTGCT 2220
Qy 2221 CAGCTCGAGGCGCTGCTTCTAGAACCTGCTGACCTGCTGCTGCTGCTGCTGCTGCT 2280
Db 2221 CAGCTCGAGGCGCTGCTTCTAGAACCTGCTGACCTGCTGCTGCTGCTGCTGCTGCT 2280
Qy 2281 TTGTCGCCAGAACAGTGGGCTGGGGAGAACTGAGACCTTGTCTGGGGGAGCATCGAGC 2340
Db 2281 TTGTCGCCAGAACAGTGGGCTGGGGAGAACTGAGACCTTGTCTGGGGGAGCATCGAGC 2340
Qy 2341 ACCACTCGCCATGTGAGCATGTTTTTCTCAATGACCTTGTCAATGGCACCCCTGTC 2400
Db 2341 ACCACTCGCCATGTGAGCATGTTTTTCTCAATGACCTTGTCAATGGCACCCCTGTC 2400
Qy 2401 AAACGAGGTGGAGAGAGCTGCTCATCAGTGGGAGAACTGAGACCTTGTCTGGGGGAGCATCGAGC 2460
Db 2401 AAACGAGGTGGAGAGAGCTGCTCATCAGTGGGAGAACTGAGACCTTGTCTGGGGGAGCATCGAGC 2460
Qy 2461 AGACGAGTGTGATGAGAGAGCAATCTGCACTGGGAGCACTTTTGCTCCAATGGA 2520
Db 2461 AGACGAGTGTGATGAGAGAGCAATCTGCACTGGGAGCACTTTTGCTCCAATGGA 2520
Qy 2521 TCATTCATGCTATTCAGAGCTTTAAAGACACCCCTGCGCCCTGACCTTGAACCCACA 2580
Db 2521 TCATTCATGCTATTCAGAGCTTTAAAGACACCCCTGCGCCCTGACCTTGAACCCACA 2580
Qy 2581 CTGGCTCGGAGCAGGATCTCTTTCGATCACTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Db 2581 CTGGCTCGGAGCAGGATCTCTTTCGATCACTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Qy 2641 CTGCTTACTCTACTTTTCGATCAACCGAGAGCAATCGGCTTCCAGCAATTTGAGTGG 2700
Db 2641 CTGCTTACTCTACTTTTCGATCAACCGAGAGCAATCGGCTTCCAGCAATTTGAGTGG 2700
Qy 2701 AAGAGGACATTAATGTTGAGCTCTTTCGATCAACCGAGAGCAATCGGCTTCCAGCAATTTGAGTGG 2760
Db 2701 AAGAGGACATTAATGTTGAGCTCTTTCGATCAACCGAGAGCAATCGGCTTCCAGCAATTTGAGTGG 2760
Qy 2761 TGATGAGCAGACACCTCAGCTCCCGAGAACCTTCTGAGACCCCTTCAAGCACTG 2820
Db 2761 TGATGAGCAGACACCTCAGCTCCCGAGAACCTTCTGAGACCCCTTCAAGCACTG 2820
Qy 2821 AAGAACGCGAGCTTGGAGGCAATGACCCCTTGGAGACCTGAGGAGCTGAGGAGGAG 2880
Db 2821 AAGAACGCGAGCTTGGAGGCAATGACCCCTTGGAGACCTGAGGAGCTGAGGAGGAG 2880
Qy 2881 TGCCAGCCATCAGTCACTGCGCACTTGGGCAATCACTGAGATTTCTGAGCACTGAGTGGC 2940
Db 2881 TGCCAGCCATCAGTCACTGCGCACTTGGGCAATCACTGAGATTTCTGAGCACTGAGTGGC 2940
Qy 2941 TTTTAGGAACCTAAGTCTTTAGCACTGAGAGCCATCACTCATCTCTCTGCTGATC 3000
Db 2941 TTTTAGGAACCTAAGTCTTTAGCACTGAGAGCCATCACTCATCTCTCTGCTGATC 3000
Qy 3001 TGGGGGTGTTTCTGCTGGGTGAGAGATGTTGCTGTGCTGCCACCCAGTACAGCTTCTCC 3060
Db 3001 TGGGGGTGTTTCTGCTGGGTGAGAGATGTTGCTGTGCTGCCACCCAGTACAGCTTCTCC 3060

Qy 3061 TCTGACCCCTTGGCTCTTCTCTCTTGTACTCTTCTGAGTGGACCTGCTCCATTCGCC 3120
Db 3061 TCTGACCCCTTGGCTCTTCTCTCTTGTACTCTTCTGAGTGGACCTGCTCCATTCGCC 3120
Qy 3121 TACATGATGGGTAACTGTGATCTTCTTCTCTGTTAGATTGTAAGCCTCCTCTTGTAT 3180
Db 3121 TACATGATGGGTAACTGTGATCTTCTTCTCTGTTAGATTGTAAGCCTCCTCTTGTAT 3180
Qy 3181 CCCAGCCCTAGCCAGTGCCTGACACAGGAACTGTGCAATAAAGGTTTATGGAACAG 3240
Db 3181 CCCAGCCCTAGCCAGTGCCTGACACAGGAACTGTGCAATAAAGGTTTATGGAACAG 3240
Qy 3241 AAAAAAAAAAAAAAAAAAAAAA 3260
Db 3241 AAAAAAAAAAAAAAAAAAAAAA 3260

RESULT 2
AF160476
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 3260
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
171..2864
/codon_start=1
/product="CD44-like precursor FELL"
/protein_id="AAF82398.1"
/translation="MPQVLRVHVACHOLLENLKLISNATSLQGEPIVSVSOSTVY
INNAKIISSDIISTNGIVHIDKLSPKLLITPKNSGRLLQNLTLTNNGYIKF
SNLIQDSGLSVITDPHTPTVLPWPDQALHALPAEQDQFLNQDNKKLKYLKEH
VIRDAKVALDPTSTAWKTQSGELSVKCGAGRDIGDLFNGQCTCRIVORELLFDLG
VAGIDCLLDPTLGGRCDDTFTFDASGECGSCVNTSPSPKPKGKQKLVNLPF
KRLNEGRCERSLVIOIPRCCKGYFGDGCQPCGPDAPCNRRGVCLDQTSATCEKC
NTGFGTACMCPWFGPDLPCGCGSDHCGDDGTGSGOCLCETGWTGSCDQAV
LPVCTPCSAHATCKENNTCECNLDYEGDITCTVVDVDFQDNGGCAKVARCSQGT
KVSCCKGKGDGSHCTEIDPCADGLNGGCHHATCKMTGPGKHCKECKSHYVGDGL
NCSPQLPDRICQDNGCHADKCDVLDHFDTTGVFHLRSLPGLQKLTLPDKAREAC
ANEAATWVFNQLSYAKAKYHLCSAGWLETGRVAYPTAFASQCGSGVGVIVDYGPR
PNKSEMDVFCYRMKDVNCTCKVYVGDGFCGSLNLQVLMSPSLNFTNTEVLAYN
SSARGAFLBHLTDLISIRGTLFVPSNGLGENETLSGRDIEHLHNVSMFVNDLVNG
TTLQTRVSKLLITASQDLPQPTETRFVDGRAILQWDIFASGHLIHLVSLPKAPPAP
VTLTHGLGAGIFFAILVTGAVALAAYSYFRINRRRTIGFQHFSEEDINVAALGKQO
PENISNPLYESTTSAPPEPSYDPTDSEERQLEGNDLRLT"

CDS

171..2864
/codon_start=1
/product="CD44-like precursor FELL"
/protein_id="AAF82398.1"
/translation="MPQVLRVHVACHOLLENLKLISNATSLQGEPIVSVSOSTVY
INNAKIISSDIISTNGIVHIDKLSPKLLITPKNSGRLLQNLTLTNNGYIKF
SNLIQDSGLSVITDPHTPTVLPWPDQALHALPAEQDQFLNQDNKKLKYLKEH
VIRDAKVALDPTSTAWKTQSGELSVKCGAGRDIGDLFNGQCTCRIVORELLFDLG
VAGIDCLLDPTLGGRCDDTFTFDASGECGSCVNTSPSPKPKGKQKLVNLPF
KRLNEGRCERSLVIOIPRCCKGYFGDGCQPCGPDAPCNRRGVCLDQTSATCEKC
NTGFGTACMCPWFGPDLPCGCGSDHCGDDGTGSGOCLCETGWTGSCDQAV
LPVCTPCSAHATCKENNTCECNLDYEGDITCTVVDVDFQDNGGCAKVARCSQGT
KVSCCKGKGDGSHCTEIDPCADGLNGGCHHATCKMTGPGKHCKECKSHYVGDGL
NCSPQLPDRICQDNGCHADKCDVLDHFDTTGVFHLRSLPGLQKLTLPDKAREAC
ANEAATWVFNQLSYAKAKYHLCSAGWLETGRVAYPTAFASQCGSGVGVIVDYGPR
PNKSEMDVFCYRMKDVNCTCKVYVGDGFCGSLNLQVLMSPSLNFTNTEVLAYN
SSARGAFLBHLTDLISIRGTLFVPSNGLGENETLSGRDIEHLHNVSMFVNDLVNG
TTLQTRVSKLLITASQDLPQPTETRFVDGRAILQWDIFASGHLIHLVSLPKAPPAP
VTLTHGLGAGIFFAILVTGAVALAAYSYFRINRRRTIGFQHFSEEDINVAALGKQO
PENISNPLYESTTSAPPEPSYDPTDSEERQLEGNDLRLT"

ORIGIN

Query Match 100.0%; Score 3259; DB 9; Length 3260;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGCACGACGAGGAGCTTCCCAAGAACCCCAAACTCCCAAGTATTTCTCCAGTTCAGG 60

Db	1	GGCAGAGAGAGCTTCCAGAACCCGAAACCTTCCAGTATTTCTTCCAGTTGCAGG	60	1081	ACTGTGAGGCTCCCTGGAGGACAGATGCCCCGTGTAAATAACCGGGGTGTCTGCTTG	1140
Qy	61	AGCATTTTGGTAAAGATCTGGTGGCCAGGCCCTTCACTTTTGGTGCACCTTATCTG	120	1141	ATCAGTACTTCGGCCACCGGAGAGTCTAAATGCAACACCGCTTCAATGGGACGGGTG	1200
Db	61	AGCATTTTGGTAAAGATCTGGTGGCCAGGCCCTTCACTTTTGGTGCACCTTATCTG	120	1141	ATCAGTACTTCGGCCACCGGAGAGTCTAAATGCAACACCGCTTCAATGGGACGGGTG	1200
Qy	121	CAGCCTTTGATGAGGAAGCTCGGTTAAAGATCTGGGACAAATACGGTTTAAATGCCCCAG	180	1201	AGATGTGTGGCCGGGAGATTTCGGGCTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTG	1260
Db	121	CAGCCTTTGATGAGGAAGCTCGGTTAAAGATCTGGGACAAATACGGTTTAAATGCCCCAG	180	1201	AGATGTGTGGCCGGGAGATTTCGGGCTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTG	1260
Qy	181	TTCTTCGGTACCATGTGGTGGCTGCCACCGCTGCTTCTGAAAACCTGAAATGATCT	240	1261	GACATGGGATGATGGCATCAGGGCTCGGGCAGTGCCTCTGTGAAACGGGGTGACAG	1320
Db	181	TTCTTCGGTACCATGTGGTGGCTGCCACCGCTGCTTCTGAAAACCTGAAATGATCT	240	1261	GACATGGGATGATGGCATCAGGGCTCGGGCAGTGCCTCTGTGAAACGGGGTGACAG	1320
Qy	241	CAAAATGCTACTTCCCTCAAGAGAGCCAAATAGTATCTCCGTCTCTCAGAGACGGTGT	300	1321	GGCCCTCGTGTGACACTCAGGCGATTTTCCCTGCGTGTGATGCTGCTGCTGCTGCTG	1380
Db	241	CAAAATGCTACTTCCCTCAAGAGAGCCAAATAGTATCTCCGTCTCTCAGAGACGGTGT	300	1321	GGCCCTCGTGTGACACTCAGGCGATTTTCCCTGCGTGTGATGCTGCTGCTGCTGCTG	1380
Qy	301	ATATAAATTAAGGCTAAGATCATATCCAGTGATATCATAGTACTAATGGATTGTTTC	360	1381	ATGCCACCTGTAAAGAGAAACAAACGCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTG	1440
Db	301	ATATAAATTAAGGCTAAGATCATATCCAGTGATATCATAGTACTAATGGATTGTTTC	360	1381	ATGCCACCTGTAAAGAGAAACAAACGCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTG	1440
Qy	361	ATATCATAGCAAAATGCTATCTCCCAAAATTTGCTTATCACTCCCAAGAGCAACTCTG	420	1441	TCACATGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
Db	361	ATATCATAGCAAAATGCTATCTCCCAAAATTTGCTTATCACTCCCAAGAGCAACTCTG	420	1441	TCACATGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1500
Qy	421	GAAGAAATCTGCAAAATCTTACGATTTTGGCAACAAACAAATGCTACATCAAAATTTAGCA	480	1501	GATGTCTCCAGAGGAGGACGAAAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1560
Db	421	GAAGAAATCTGCAAAATCTTACGATTTTGGCAACAAACAAATGCTACATCAAAATTTAGCA	480	1501	GATGTCTCCAGAGGAGGACGAAAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1560
Qy	481	ACTTATACAGGACTCAGTTTGTCTGAGTGTATCATCCGATCCATCCACACCCAGTCA	540	1561	GGCAGAGCTGACAGAGATAGACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620
Db	481	ACTTATACAGGACTCAGTTTGTCTGAGTGTATCATCCGATCCATCCACACCCAGTCA	540	1561	GGCAGAGCTGACAGAGATAGACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620
Qy	541	CTCTCTTCTGGCCACCGACCAAGCCCTCCATGCCCTTACCTGCTGCAACACAGGACTTC	600	1621	AGCCACCTGTAAAGATGACAGGCGGGGCAAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTG	1680
Db	541	CTCTCTTCTGGCCACCGACCAAGCCCTCCATGCCCTTACCTGCTGCAACACAGGACTTC	600	1621	AGCCACCTGTAAAGATGACAGGCGGGGCAAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTG	1680
Qy	601	TGTTCAACCAAGACAAAGCAAGCTGAAGAGTATTTGAAGTTTCTGATACGAG	660	1681	TGGAGATGGCTGAACTGTGAGCGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1740
Db	601	TGTTCAACCAAGACAAAGCAAGCTGAAGAGTATTTGAAGTTTCTGATACGAG	660	1681	TGGAGATGGCTGAACTGTGAGCGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1740
Qy	661	ATGCCAAGTTTGTAGCTGTGATCTTCCACATCCATGCTGCTGCAAGACCCCTCAAGTT	720	1741	ATGGGAGTGCCTATGACAGCGCCAAATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800
Db	661	ATGCCAAGTTTGTAGCTGTGATCTTCCACATCCATGCTGCTGCAAGACCCCTCAAGTT	720	1741	ATGGGAGTGCCTATGACAGCGCCAAATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1800
Qy	721	CAGAGCTGAGTGTGAATGTGGAGCTGGCAGGAGCATCGGTGACCTCTTCTGAATGGCC	780	1801	GGGTGTTCATCTACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1860
Db	721	CAGAGCTGAGTGTGAATGTGGAGCTGGCAGGAGCATCGGTGACCTCTTCTGAATGGCC	780	1801	GGGTGTTCATCTACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1860
Qy	781	AAACCTGAGAAATGTGAGGGAGCTCTTGTGACCTGGGTGGCTTACGGCATTG	840	1861	AGCCCTGTGCAACGAAGCTGCGACCATCTGCAACCTTACCAACCTGCTGCTGCTGCTGCTG	1920
Db	781	AAACCTGAGAAATGTGAGGGAGCTCTTGTGACCTGGGTGGCTTACGGCATTG	840	1861	AGCCCTGTGCAACGAAGCTGCGACCATCTGCAACCTTACCAACCTGCTGCTGCTGCTGCTG	1920
Qy	841	ACTGTCTGCTGATTTGATCCACCTGGGGGCGCTGTGACACTTACTTCTGATG	900	1921	AGCCAAAGTACCACTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1980
Db	841	ACTGTCTGCTGATTTGATCCACCTGGGGGCGCTGTGACACTTACTTCTGATG	900	1921	AGCCAAAGTACCACTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1980
Qy	901	CCTCGGGGAGTGTGGAGCTGTGTAATACTCCAGCTGCCCAAGTGGAGTAAACCAA	960	1981	CAGCCTTCCCTCCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2040
Db	901	CCTCGGGGAGTGTGGAGCTGTGTAATACTCCAGCTGCCCAAGTGGAGTAAACCAA	960	1981	CAGCCTTCCCTCCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2040
Qy	961	AGGTGTGAGCAGAGTGTCTTACAACTGCTTCAAGAGGAACTGGAAGGCTGCC	1020	2041	GACCCAAAGAGTGAATGTGGATGTCTTCTGCTATCGGATGAAAGATGTGAATGCA	2100
Db	961	AGGTGTGAGCAGAGTGTCTTACAACTGCTTCAAGAGGAACTGGAAGGCTGCC	1020	2041	GACCCAAAGAGTGAATGTGGATGTCTTCTGCTATCGGATGAAAGATGTGAATGCA	2100
Qy	1021	GGAGCGGTGCACTGCTGATACAGATCCCGAGTCTGCAAGGGTACTTTCGGGCGAG	1080	2101	CCTGCAAGGTGGGTATGTGGAGATGGCTTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2160
Db	1021	GGAGCGGTGCACTGCTGATACAGATCCCGAGTCTGCAAGGGTACTTTCGGGCGAG	1080	2101	CCTGCAAGGTGGGTATGTGGAGATGGCTTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2160
Qy	1081	ACTGTACGGCTGCTCGGAGGACAGATGCCCGTGTAAATAACCGGGGTGTCTGCTTG	1140	2161	TGATGTCTTCTTCCCTTACTCACAACCTTCTCTGACGGAAGTGTGGCTTATTCACACAGCT	2220
Db	1081	ACTGTACGGCTGCTCGGAGGACAGATGCCCGTGTAAATAACCGGGGTGTCTGCTTG	1140	2161	TGATGTCTTCTTCCCTTACTCACAACCTTCTCTGACGGAAGTGTGGCTTATTCACACAGCT	2220

KVLEIQKRNCDNNAITLIRRCRITCSSELTCPGFKSLGNKRRRCIYTSYPMGRRLTF
IGCQPKVTVLITRECCAGFPQCPQCPGNAQNVCFNGICLDGVNGTGVCEGEGE
SGTACETEGKGLHCDQACSCVHGRNQNGELGDSGCDVGMGRVGHCDNAITEDNC
NGTCHTSANLNSGTSKCAAGFQNGGTCTAINACEISNGSGSADAKCRITPG
RRVCTKAGYTDGGLVLEINFLENHGGCDKNAECTQTPNQAAACNCLPAYTGBGK
CTLINVTIKNGGCEFAFICNHTGOVETCTCKPNYIGDGFTRGSIYOELEPNPKTS
QYFPOLEHFDVLDVGPFTVFAPLSAFADEARVKDWDKYLMPQVLYHVHACHQ
LLELNLISNATSGOEPILVSOSTVYINNAKILSSDIISTNGIWHIIDKLLSP
KNLLTIPKNSGRILONLTLATNNGYIKFNSLIQDSGLLSVITDPIHTPVLFPPTD
QALHALPASQDPLFMDNKKLKEYLKEFHVIRDAKVLAVDLPTSTAWKTLOSSELSV
KCGAGRIDGLFLNGQTCRIVQRELLFDLGVAYGIDCLLIDPTLGGRCIDFTTFDASG
EGSCVTPSCSPKSPKGVOKLYNLFKNLEGRERCESLVIQIPRCCKGYFRD
COACPGPDPCNNRVCCLDQYATCEKCNTEFNAGCEMCPWPRGFDCLPCGCSG
HGOCDGLTSGOCLCETCWTPGSDTOAVLPAVCTPPCSAHATCKENNTCECNLDYE
GGITCTVVDKQDNGGCAKVARCSQKTKVSCQKYGKGDHSCHEIDPCADGLN
GGCHEATKMTGPKHKCEKSHYVGDGLNCEPEQLPIDRCLQDNGQCHADAKCVDL
LFQDITVYFHLRSLPGQYKLTFRKAREACANEATMYNQLSYAKAKYHLCISAGW
HEDTRVATPTAFASQSGSVGIVDYGRPNKSEMDWVFCYRMKDINVTCKVYQSG
GFSCSNLILQVMSFPLTNFLTEVLAYSNSARGRAFLHLETLISIRGLTFVPOKSG
LGENETLSGRDIEHHLANVSMEFYNDLVNLTLOTLRGLSKLLITASQDPIOTETRFV
DGRALLOWIDFASNGLIHVISRPLKAPPAVTLTHGLGAGIFPAILLVGTGAVALAY
SIFRINRRITGFQHFSEEDINVAALGKQOOPENISNPLYESTISAPPPEPSYDPTDSE
BRLEGNDPRLT"

ORIGIN

	Query Match	Score	3241.4;	DB 9;	Length	8266;
	Best Local Similarity	99.7%;	Pred. No. 0;			
	Matches 3245;	Conservative	1;	Mismatches	8;	Indels 0; Gaps 0;
QY	7	AGCAGAGCTTCCCAAGAACCCGAAACTTCCACGATTTCTTCCAGTTTGCAGAGCAT	66			
Db	4985	ATCAGAGCTTCCCAAGAACCCGAAACTTCCACGATTTCTTCCAGTTTGCAGAGCAT	5044			
QY	67	TGCTGAAGATCTGGTGGCCGAGCCCTTCACTGTTTTCACCTTTATCTGAGCCT	126			
Db	5045	TGCTGAAGATCTGGTGGCCGAGCCCTTCACTGTTTTCACCTTTATCTGAGCCT	5104			
QY	127	TTGATCAGAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCGAGTCTTTC	186			
Db	5105	TTGATCAGAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCGAGTCTTTC	5164			
QY	187	GGTACCATGTGGTGGCTGCCACAGCTGTTCTCGAAACCTGAAATGATCTCAATG	246			
Db	5165	GGTACCATGTGGTGGCTGCCACAGCTGTTCTCGAAACCTGAAATGATCTCAATG	5224			
QY	247	CTACTTCCCTCAAGAGAGCCAAATAGTCACTCCGCTCTCTCAGAGCAAGGTATATAA	306			
Db	5225	CTACTTCCCTCAAGAGAGCCAAATAGTCACTCCGCTCTCTCAGAGCAAGGTATATAA	5284			
QY	307	ATAATAAGGCTAAGATCATATCAGTGATATCATAGTACTAATGGATTTGTCATATCA	366			
Db	5285	ATAATAAGGCTAAGATCATATCAGTGATATCATAGTACTAATGGATTTGTCATATCA	5344			
QY	367	TAGACAAATGTCTATCCCAAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGAA	426			
Db	5345	TAGACAAATGTCTATCCCAAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGAA	5404			
QY	427	TTCTGCAAAATCTTAGACTTTGGCAACAAAGATGGCTACATCAAAATTTAGCAACTTAA	486			
Db	5405	TTCTGCAAAATCTTAGACTTTGGCAACAAAGATGGCTACATCAAAATTTAGCAACTTAA	5464			
QY	487	TACAGACTCAGGTTTGTGAGTGTATCACCGATCCCATCCACACCCAGTCACTCTCT	546			
Db	5465	TACAGACTCAGGTTTGTGAGTGTATCACCGATCCCATCCACACCCAGTCACTCTCT	5524			
QY	547	TCTGGCCCAACGACCAAGCCCTCATGCCCTACTCCTGTGAACAAACAGGACTTCCGTGCA	606			
Db	5525	TCTGGCCCAACGACCAAGCCCTCATGCCCTACTCCTGTGAACAAACAGGACTTCCGTGCA	5584			
QY	607	ACCAAGACAAAGGACAAAGCTCAAGAGTATTTGAAGTTTCATGTGATACAGATGCCA	666			
Db	5585	ACCAAGACAAAGGACAAAGCTCAAGAGTATTTGAAGTTTCATGTGATACAGATGCCA	5644			

QY	667	AGGTTTGTAGCTGTGAGATCTTCCACATCCATGCTGCTGGAAGACCCCTGCAAGGTTTCAAGC	726
Db	5645	AGGTTTGTAGCTGTGAGATCTTCCACATCCATGCTGCTGGAAGACCCCTGCAAGGTTTCAAGC	5704
QY	727	TCAGTGTGAAATGTGGAGCTGGCAGGGAATCGGTGACCTCTTTTCTGAATGGCCAAACCT	786
Db	5705	TCAGTGTGAAATGTGGAGCTGGCAGGGAATCGGTGACCTCTTTTCTGAATGGCCAAACCT	5764
QY	787	GCAGAAATTTGTCAGCGGGAGCTCTTGTGACCTGGTCTGGCTTACGCAATGACTGTTC	846
Db	5765	GCAGAAATTTGTCAGCGGGAGCTCTTGTGACCTGGTCTGGCTTACGCAATGACTGTTC	5824
QY	847	TGCTGATTCATCCCAACCTCGGGGGCCGTGTGACACACCTTTACTTTCTGATCGCTCGG	906
Db	5825	TGCTGATTCATCCCAACCTCGGGGGCCGTGTGACACACCTTTACTACTTTTCTGATCGCTCGG	5884
QY	907	GGAGTGTGGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG	966
Db	5885	GGAGTGTGGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG	5944
QY	967	TGAAGCAGAGTGTCTCTACACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC	1026
Db	5945	TGAAGCAGAGTGTCTCTACACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGGAGC	6004
QY	1027	GGTGCAGCTGTGTATACAGATCCCAGGTGTCTGCAAGGGCTACTTTCCGGCCGAGACTGTTC	1086
Db	6005	GGTGCAGCTGTGTATACAGATCCCAGGTGTCTGCAAGGGCTACTTTCCGGCCGAGACTGTTC	6064
QY	1087	AGGCTGCTGCGAGGACAGATGCCCCGTGTAAATCAACCGGGGTGTCTGCTTGTGATCAGT	1146
Db	6065	AGGCTGCTGCGAGGACAGATGCCCCGTGTAAATCAACCGGGGTGTCTGCTTGTGATCAGT	6124
QY	1147	ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGGTGTGAGATGT	1206
Db	6125	ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGGTGTGAGATGT	6184
QY	1207	GCTGGCCGGGAGATTCGGGCTCTGATTTGTCTGCTGCTGCTGCTCAGACACCGGACAGT	1266
Db	6185	GCTGGCCGGGAGATTTGGGCTCTGATTTGTCTGCTGCTGCTCAGACACCGGACAGT	6244
QY	1267	GGATGATGGCATCAGGGCTCCGGGCTCTGTGTGAAACCGGGGTGACAGGCCCT	1326
Db	6245	GGATGATGGCATCAGGGCTCCGGGCTCTGTGTGAAACCGGGGTGACAGGCCCT	6304
QY	1327	CGTGTGACACTCAGGAGTTTTGCCGTGAGTGTACGCTCTCTCTGCTCTCATGCCA	1386
Db	6305	CGTGTGACACTCAGGAGTTTTGCCGTGAGTGTACGCTCTCTCTGCTCTCATGCCA	6364
QY	1387	CCTGTAAGGAGAACACACGTGTGATTAACCTGGATTTATGAAGTGAACGGAATCACAT	1446
Db	6365	CCTGTAAGGAGAACACACGTGTGATTAACCTGGATTTATGAAGTGAACGGAATCACAT	6424
QY	1447	GCACAGTTGTGATTTCTGCAACAGGACAAACCGGGCTGTGCAAGGTTGGCCAGATGCT	1506
Db	6425	GCACAGTTGTGATTTCTGCAACAGGACAAACCGGGCTGTGCAAGGTTGGCCAGATGCT	6484
QY	1507	CCAGAAAGGACAGAAAGTCTCTCTGAGCTGCCAGAGGATACAAAGGGGACGGGACCA	1566
Db	6485	CCAGAAAGGACAGAAAGTCTCTCTGAGCTGCCAGAGGATACAAAGGGGACGGGACCA	6544
QY	1567	GCTGCAACAGATAGACCCCTGTGACAGCGGCTTAAACGGAGGTTGTACGAGCACGCCA	1626
Db	6545	GCTGCAACAGATAGACCCCTGTGACAGCGGCTTAAACGGAGGTTGTACGAGCACGCCA	6604
QY	1627	CCTGTAAGATGACAGGCGGGCAAGCAAGTGTGAGTGAAGTCTACTATGTCGGAG	1686
Db	6605	CCTGTAAGATGACAGGCGGGCAAGCAAGTGTGAGTGAAGTCTACTATGTCGGAG	6664
QY	1687	ATGGGCTGAACTGTGAGCGGGAGCAGCTGCCATTCACCGCTGTTTACAGGACAAATGGGC	1746
Db	6665	ATGGGCTGAACTGTGAGCGGGAGCAGCTGCCATTCACCGCTGTTTACAGGACAAATGGGC	6724
QY	1747	AGTGCCATGCAGACGCCAAATGTGTGACCTCCACTTCCAGGATACCACTGTTGGGTGT	1806


```
Db 6725 AGTGCATGACAGCCCAAAATGTGCGACCTCACTTCAGGATACCACTGTGGGGTGT 6784
QY 1807 TCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTTACAAAGCCAGAGAGGCT 1866
Db 6785 TCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTTACAAAGCCAGAGAGGCT 6844
QY 1867 GTGCCAAGAGCTGGACCATGGCAACTCAACACAGCTTCCTATGCCCCAGAGAGCCA 1926
Db 6845 GTGCCAAGAGCTGGACCATGGCAACTCAACACAGCTTCCTATGCCCCAGAGAGCCA 6904
QY 1927 AGTACCACTGTGCTCAGCAGCTGCTGGAGACCGGGGGTTCCTACCCACAGCT 1986
Db 6905 AGTACCACTGTGCTCAGCAGCTGCTGGAGACCGGGGGTTCCTACCCACAGCT 6964
QY 1987 TCGCCTCCAGAACTGTGGCTCTGGTGTGGATAGTGAGCTATGGACCTAGACCCA 2046
Db 6965 TCGCCTCCAGAACTGTGGCTCTGGTGTGGATAGTGAGCTATGGACCTAGACCCA 7024
QY 2047 ACAAGAGTGAATGTGGAGTGTCTCTGCTATCGGATGAAGATGTGAACCTGGA 2106
Db 7025 ACAAGAGTGAATGTGGAGTGTCTCTGCTATCGGATGAAGATGTGAACCTGGA 7084
QY 2107 AGTGGGCTATGTGGAGATGCTTCTCATGAGTGGAACTGCTGCAGGTTCCTGATGT 2166
Db 7085 AGTGGGCTATGTGGAGATGCTTCTCATGAGTGGAACTGCTGCAGGTTCCTGATGT 7144
QY 2167 CTTTCCCTCACTCAAACTTCTCTGCGGAGTGTCTGCTATTCACAGCTCAGCTC 2226
Db 7145 CTTTCCCTCACTCAAACTTCTCTGCGGAGTGTCTGCTATTCACAGCTCAGCTC 7204
QY 2227 GAGGCGTGCATTTCTAGAACCTGACTGACCTGTCCATCGCGGCACTCTTTTGTGC 2286
Db 7205 GAGGCGTGCATTTCTAGAACCTGACTGACCTGTCCATCGCGGCACTCTTTTGTGC 7264
QY 2287 CACGAACAGTGGCTGGGGAGATGAGACTTCTGTGGGGGACATCGAGCACACC 2346
Db 7265 CACGAACAGTGGCTGGGGAGATGAGACTTCTGTGGGGGACATCGAGCACACC 7324
QY 2347 TCGCAATGTGACGATGTTTTTCTAATGACCTTGTCAATGGCACCCCTCGAAACGA 2406
Db 7325 TCGCAATGTGACGATGTTTTTCTAATGACCTTGTCAATGGCACCCCTCGAAACGA 7384
QY 2407 GGTGGGAAGCAGTGTCTCATCTGCGAGCAGGACCCACTCAACCGAGCGAGACA 2466
Db 7385 GGTGGGAAGCAGTGTCTCATCTGCGAGCAGGACCCACTCAACCGAGCGAGACA 7444
QY 2467 GGTGTGTGATGGAAGAGCCATCTGCGAGTGGACATCTTTCCTCCCAATGGGATCTC 2526
Db 7445 GGTGTGTGATGGAAGAGCCATCTGCGAGTGGACATCTTTCCTCCCAATGGGATCTC 7504
QY 2527 ATGTCAATTCAGGCTTTTAAAGCACCCCTGTCGCCCGGTGACCTTGACCCACCTGGCT 2586
Db 7505 ATGTCAATTCAGGCTTTTAAAGCACCCCTGTCGCCCGGTGACCTTGACCCACCTGGCT 7564
QY 2587 TGGGAGCAGGATCTTCTTTGCGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 2646
Db 7565 TGGGAGCAGGATCTTCTTTGCGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 7624
QY 2647 ACTCCTACTTTTGGGATAAAGCGGAAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAG 2706
Db 7625 ACTCCTACTTTTGGGATAAAGCGGAAACAATCGGCTTCCAGCATTTTGAGTCGGAAGAG 7684
QY 2707 ACATTAATTTGACGCTCTTGSCAAGCAGCAGCTGAGAATATCTCGAACCCCTTGATG 2766
Db 7685 ACATTAATTTGACGCTCTTGSCAAGCAGCAGCTGAGAATATCTCGAACCCCTTGATG 7744
QY 2767 AGAGCACAACCTCAGCTCCCCAGAACCTTCTCAGACCCCTTCAGGACTCTGAAGAAC 2826
Db 7745 AGAGCACAACCTCAGCTCCCCAGAACCTTCTCAGACCCCTTCAGGACTCTGAAGAAC 7804
QY 2827 GGCAGCTTGAGGGCAATACCCCTTTGAGGACACTGTGAGGGCTCGAGCGGAGATGCCAG 2886
```

```
Db 7805 GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCTGGACGGGAGATGCCAG 7864
QY 2887 CCATCACTCACTGCCACCTGGGCCCATCAACTGTGAATTTCTCAGCACCCAGTTGCCCTTTT 2946
Db 7865 CCATCACTCACTGCCACCTGGGCCCATCAACTGTGAATTTCTCAGCACCCAGTTGCCCTTTT 7924
QY 2947 GAACCTAAAGTCTTTTAAAGCACTCAGAGCCATCCTCATCTCTGCTGCTGATCTGGGG 3006
Db 7925 GAACCTAAAGTCTTTTAAAGCACTCAGAGCCATCCTCATCTCTGCTGCTGATCTGGGG 7984
QY 3007 TTGTTTCTGCTGGTGAGAGATGTGCTGTGCCACCCAGTACAGCTTCTCCTCTGAC 3066
Db 7985 TTGTTTCTGCTGGTGAGAGATGTGCTGTGCCACCCAGTACAGCTTCTCCTCTGAC 8044
QY 3067 CTTTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3126
Db 8045 CTTTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 8104
QY 3127 ATGGGTAAGTGTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3186
Db 8105 ATGGGTAAGTGTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8164
QY 3187 CCCTAGCCCACTGCTGACACAGGAACCTGTCACAATAAAGTTTATGGAACAGAAAAA 3246
Db 8165 CCCTAGCCCACTGCTGACACAGGAACCTGTCACAATAAAGTTTATGGAACAGAAAAA 8224
QY 3247 AAAAAAAGAAAAA 3260
Db 8225 AGTCAAAAAA 8238
```

```
RESULT 4
AY311388
LOCUS Homo sapiens FEX2 mRNA, complete cds. 8251 bp mRNA linear PRI 30-JUN-2003
DEFINITION
ACCESSION AY311388
VERSION AY311388.1 GI:32351284
```

```
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 8251)
AUTHORS Park,S.-Y. and Kim,I.-S.
TITLE FEX2, a novel cell adhesion molecule of Fas-1 superfamily mediates
cell-cell interaction
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 8251)
AUTHORS Park,S.-Y. and Kim,I.-S.
TITLE Direct Submission
JOURNAL
MEDICINE, Kyungpook National University, 101 Dongin-dong, Jung-gu,
Deagu 700-422, Korea
```

```
FEATURES
Location/Qualifiers
1..8251
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q23.3"
205..7860
/note="FAS1, EGF-like and X-link domain containing
adhesion molecule-2"
```

```
CDS
translation="MMLQLVIFCLGVQVNFCSPTTGOARRCDRLSLITRECR
SCALNGLVKPDGTYMTITSGVGRDCTYFEVRYTSLSPGCRHI CRKDYLRCCP
GRWPDCTIECPGAGSPCNGRCSACBGMGNTCSQEGFGTACETACDNLFGPSC
SSVPCNVHVCNSGLDGTCECYSAVTGPKCDKPIPECCALLCPENSRCSPTFEDN
KLECKLPLNRYGDKYCDPINPLRKIKIHPHAKTYLIGPNRHSCTCEGYRGDQVCL
PVDPCQINFNCPTKSVCKYDGGQSCHECKEYQNFVPGVGCSTMDICKSDNPFCH
```

NANCTTVAQRTECICOKGYVGDGLTCYGNIMERLRELNTEPRGKWOQRLTSTFISILDD
 KAYANPLSKLGFPTVLLPTDCKLKGPNVNEILLVDNKAQYFVKLHLIAGOMNIEYWN
 TDMFTLTKSGRIITFNSDKONQIKLKHGGKKVKLIQGDIIASNGLLHILDRADMLK
 EPTFESNEQTIMLQOPRYSKRSLLEETNEHALDEDEGVGPVYI FVPNNEALNM
 KOGTLDYLLSPGSKLLELHYHIVPFTQLEVAITLSTPHSMAQULIOFNITDNG
 QILANDVAMEETITAKNGRIYTLGVLIPPSIPLPHRCDETKREMLKGTVCSSL
 VYWSRCPANSEPTALFTHRCVYSGRFLSKSGCARVNAIVKI PKCKGFFYPCNQOC
 PGFSNP CSGNQCADSLGGNGTCICEBGFQSQCOQFCSDPNKYGRPNCKXCLVHGT
 CNRIDSQAGTICRDSAGRLCDKQTSACGPYVQFCHIHATCEYSNGTASCICA
 GYEGDTGLSEMDTLPGLPGCSRNACIKTGTGTCVQCGQNGRDCSEINNCC
 LLPASGCHDNASCLYVGPQNECKCKGRNGIDCEPITSCLPQTKGCHPLASQCS
 TSSGVSCVOCBGEGLCYGNAALVFLSEAAIFNRWINNLSLOPTLSATNSLT
 VLVPSQOATEDMDQEKSPWLSQSNIPALIKYHMLLGTVRVADILQTLSSMDLATSLO
 GNELHLAKVDGNTIEGASIVGDNAATNGVHIINKVLVPORRLTGSLPNLLMLEQ
 MEDYSIFRGYIIIOYLANALEAADATVFPAPNNAIENITREKVLNLSLEEDVIRTHV
 LEEKLJLHNGHRETMFLPSYFLSFLPHLDQLYVNBAPI NYNTVADTKGVHGLG
 KVLBTQNRKUNDDTTI INGRCTCSELTCFPTKSLGNERRCICYTSYFMRGRTLF
 IGCQPKVYITRBCAGFGPQCPQCPGNAQNVFNGICLDGVNGTGVCECBGF
 SGTACETCTEGYKGIHDOACSCVHRCNQGLDGS CDCVGRGVHCDNATIEDNC
 NGTCHTSANCLNSDTASCKCAAGPQNGTICTAINACEIISNGGCSAKADCKRTPG
 REVCTKAGYTGDDIVCLEINPLENHGSCDKNAECTQGRPNQAACNCLPATYTGDKV
 CRLINVLTKNGCSSEFALCNHTGQVERTCTKPNYIGDFTCRGSIYOELPNFATS
 QYFFOLQBEHVDLVPGPFTVAPLSAAPDEBEARVDKWDKYLQPLVIRHVNAHQ
 LLENKLTISNATIQEPIVSVSQSTVYINNAKIIISDIIISTINGIWHIIDKLISP
 KNLLITPKDONSRILOLTLATNGYIKFNSLIODSGLLSVITDPIHPVTLFWPTD
 OALHALPAQODLFNDQNDKLEKYLKPHVIRDAKVLAVDLPSTAMKTLQGSLSV
 KCGAGRDIDGLFLNGQTCRLVORELLFLDGVAYGIDCLLIDTLGRCDDTFTTFAASG
 ECGSGVTPSCPRWSPKGVKQKCLNLPFRNLRGCRERCSLVIQIPKCKGYFDRG
 COACGPGDPANNRGVCLDQYATSGECKNTGPNGTACEMCPGRFDEPCLDPCQSD
 HQCDQDGTGQCLCETGWTGSCDIOALVPACTPCCSAHATKENNTCECNLDYE
 GDIETCTVVDQKQNGCAKVARSGKGVSKVCSQKGYKGDGHSCTEIDPCADGLN
 GGCHEHATCKMGPCKHCKSHVYVGDENCEPEOLPIDRCLDNGQCHADAKCVDL
 HFQDTTVGVFHLRSPLGOYKLPDKAREACANEATMATYNOLSVAKAKYHLCSGW
 LETGRVAYPTAFASQSGVYGVYDYGPRPNKSEMMVDFCRMDVNCITCKGVNGD
 GFSCNLLQVLMSPSLNFIETVLAVSNSARGRAFLHEHLDITLUSQDPLQTFETRFV
 LGENATLSGRDIFASNGIHIIVSRPLKAPPAPVTLTHTLGAGIFFAILIIVTGAVALAA
 SYVRINRRTIGFQHFSEEDINVAALGXQOEPENISNPLYESTTSAPPEPSYDFPTDSE
 ERQLEGNDFRLT"

ORIGIN

Query Match		99.2%;	Score 3234.2;	DB 9;	Length 8251;
Best Local Similarity		99.7%;	Pred. No. 0;		
Matches 3239;		Conservative 0;	Mismatches 9;	Indels 0;	Gaps 0;
QY	7	AGCAGAGACTTCCCAAGAACCCGAAACCTTCCAGTATTTCTTCCAGTTGCGAGGACATT	66		
DB	5003	ATCAGAGACTTCCCAAGAACCCGAAACCTTCCAGTATTTCTTCCAGTTGCGAGGACATT	5062		
QY	67	TCGTGAAGATCTGGTGGCCCGAGGCCCTTCACTGTTTTTGACCTTTATCTGCAGCCT	126		
DB	5063	TGCTGAAGATCTGGTGGCCCGAGGCCCTTCACTGTTTTTGACCTTTATCTGCAGCCT	5122		
QY	127	TTGATGAGAACTCGGGTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC	186		
DB	5123	TTGATGAGAACTCGGGTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC	5182		
QY	187	GGTACCATGTGTCGCGCCACCGACTGCTTCTGGAACCTGGAATTCATCTCAAATG	246		
DB	5183	GGTACCATGTGTCGCGCCACCGACTGCTTCTGGAACCTGGAATTCATCTCAAATG	5242		
QY	247	CTACTTCCCTCCAAAGAGACCAATAGTCAITCTCCGTCTCTCAGACACGGTGTATATAA	306		
DB	5243	CTACTTCCCTCCAAAGAGACCAATAGTCAITCTCCGTCTCTCAGACACGGTGTATATAA	5302		
QY	307	ATAATAGGCTAAGATCATATCCAGTGATATCATATAGTACTAATGGATTTGTTATATCA	366		
DB	5303	ACAATAAGGCTAAGATCATATCCAGTGATATCATATAGTACTAATGGATTTGTTATATCA	5362		
QY	367	TAGACAAATTTGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGAA	426		
DB	5363	TAGACAAATTTGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGAA	5422		

QY	427	TTCTGCARAAATCTTTAGACCTTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCAACTTAA	486
DB	5423	TTCTGCARAAATCTTTAGACCTTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCAACTTAA	5482
QY	487	TACAGACTTCAGGTTTGTCTGAGTGTCAATCACCGATCCCATCCACACCCCAAGTCACTCTCT	546
DB	5483	TACAGACTTCAGGTTTGTCTGAGTGTCAATCACCGATCCCATCCACACCCCAAGTCACTCTCT	5542
QY	547	TTCTGGCCACCGACCAAGCCCTCCATGCTTACCTGCTGTGTAACAAAGAGACTTCTCTGTTCA	606
DB	5543	TTCTGGCCACCGACCAAGCCCTCCATGCTTACCTGCTGTGTAACAAAGAGACTTCTCTGTTCA	5602
QY	607	ACCAAGACAAACAGGACCAAGCTGAAGAGTATTTTGAAGTTTCATGTGTATACAGAGTGC	666
DB	5603	ACCAAGACAAACAGGACCAAGCTGAAGAGTATTTTGAAGTTTCATGTGTATACAGAGTGC	5662
QY	667	AGGTTTATAGCTGTGATCTTCCACATCCATGCTTCCCTGGAAGACCCCTGCAAGGTTTCAGAGC	726
DB	5663	AGGTTTATAGCTGTGATCTTCCACATCCATGCTTCCCTGGAAGACCCCTGCAAGGTTTCAGAGC	5722
QY	727	TGAGTGTGAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTGAATGCGCAACCT	786
DB	5723	TGAGTGTGAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTGAATGCGCAACCT	5782
QY	787	GCAGAAATGTGCGAGCGGAGCTCTTGTGTAACCTGGGTGTGGCTACGCGCATTTGACTGTC	846
DB	5783	GCAGAAATGTGCGAGCGGAGCTCTTGTGTAACCTGGGTGTGGCTACGCGCATTTGACTGTC	5842
QY	847	TGCTGATTAATCCCAACCTTGGGGGCGCTGTGACACCTTTACTTTTCATGCTCGCTCGG	906
DB	5843	TGCTGATTAATCCCAACCTTGGGGGCGCTGTGACACCTTTACTTTTCATGCTCGCTCGG	5902
QY	907	GGGAGTGTGGAGCTGTGTCATATCTCCAGCTGCCAAGGTGGAGTAAACCAAGGGTG	966
DB	5903	GGGAGTGTGGAGCTGTGTCATATCTCCAGCTGCCAAGGTGGAGTAAACCAAGGGTG	5962
QY	967	TGAAGCAGAAAGTGTCTCTACAACTGCCCTTCAAGAGAAACCTTGAAGGCTGCGGGAGC	1026
DB	5963	TGAAGCAGAAAGTGTCTCTACAACTGCCCTTCAAGAGAAACCTTGAAGGCTGCGGGAGC	6022
QY	1027	GGTGAGCTGTGTATACAGATCCCAAGGTGTGCAAGGGCTACTTTCGGGCGAGACTGTC	1086
DB	6023	GGTGAGCTGTGTATACAGATCCCAAGGTGTGCAAGGGCTACTTTCGGGCGAGACTGTC	6082
QY	1087	AGGCTGCTCGGAGACCAAGTGCCTGTGTAACCGGGGTCTCCCTTGTGATCAGT	1146
DB	6083	AGGCTGCTCGGAGACCAAGTGCCTGTGTAACCGGGGTCTCTCCCTTGTGATCAGT	6142
QY	1147	ACTCGGCCACCGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGGTGTGAGATGT	1206
DB	6143	ACTCGGCCACCGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGGTGTGAGATGT	6202
QY	1207	GCTGGCGGGAGATTCGGGCTGTGATTTGCTGCTGTGGCTGTGCTCAGACCAAGGACAGT	1266
DB	6203	GCTGGCGGGAGATTCGGGCTGTGATTTGCTGCTGTGGCTGTGCTCAGACCAAGGACAGT	6262
QY	1267	GGGATGATGGCATACGGGCTCCGGGCTGCTGTGTAACCGGGGTGCAAGGCGCCCT	1326
DB	6263	GGGATGATGGCATACGGGCTCCGGGCTGCTGTGTAACCGGGGTGCAAGGCGCCCT	6322
QY	1327	CGTGTGACACTCAGGAGTTTTGCTGCGAGTGTGACGCTCTCTTGTCTGCTATGCA	1386
DB	6323	CGTGTGACACTCAGGAGTTTTGCTGCGAGTGTGACGCTCTCTTGTCTGCTATGCA	6382
QY	1387	CTGTGAAGAGAAACACAGCTGTGATGTAACTGTGATTTAAGGTGACGAATCACAT	1446
DB	6383	CTGTGAAGAGAAACACAGCTGTGATGTAACTGTGATTTAAGGTGACGAATCACAT	6442
QY	1447	GCACAGTTTGTGATTTCTGCAAAACAGGACCAAGGGGCTGTGCAAAAGGTGGCCAGATGCT	1506
DB	6443	GCACAGTTTGTGATTTCTGCAAAACAGGACCAAGGGGCTGTGCAAAAGGTGGCCAGATGCT	6502
QY	1507	CCCAAGAGGACGAGAGGTCTCTCTGAGCTGCCAAGAGGGATACAAAGGGGACGGGAC	1566

Db 6503 |||||CCCCGAGGACGAAAGGCTCTCCGACGCTCCGAGAGGGATACAAAGGGGACGGGACCA 6562
Qy 1567 GCTGCACAGATAGACCCCTGTGCAGACGGCTTAAACGGAGGGTGTACAGACACGCA 1626
Db 6563 GCTGCACAGATAGACCCCTGTGCAGACGGCTTAAACGGAGGGTGTACAGACACGCA 6622
Qy 1627 CCTGTAGATACAGGCCCGGGCAAGCAACAAGTGTAGTGTAAAGTCACTATGTCGGAG 1686
Db 6623 CCTGTAGATACAGGCCCGGGCAAGCAACAAGTGTAGTGTAAAGTCACTATGTCGGAG 6682
Qy 1687 ATGGGCTGAATGTGAGCGGAGCGAGCTGCCATTTGACGGCTGTACAGACATGGGC 1746
Db 6683 ATGGGCTGAATGTGAGCGGAGCGAGCTGCCATTTGACGGCTGTACAGACATGGGC 6742
Qy 1747 AGTGCCATGACAGCCCAAAATGTGTGACACTTCCAGATACCACTGTGTGGGGTGT 1806
Db 6743 AGTGCCATGACAGCGCAAAATGTGTGACACTTCCAGATACCACTGTGTGGGGTGT 6802
Qy 1807 TCCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGGCGCT 1866
Db 6803 TCCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGGCGCT 6862
Qy 1867 GTGCCAAGAGCTGCGACCAATGGCAACTACACACAGCTCTCTATGCCAGAGGCCA 1926
Db 6863 GTGCCAAGAGCTGCGACCAATGGCAACTACACACAGCTCTCTATGCCAGAGGCCA 6922
Qy 1927 AGTACCACTGTGCTCAGCAGCTGGCTGGAGACCGGGCGGTTGCCCTACCCACAGCCT 1986
Db 6923 AGTACCACTGTGCTCAGCAGCTGGCTGGAGACCGGGCGGTTGCCCTACCCACAGCCT 6982
Qy 1987 TCGCCTCCAGAACGTGGCTGTGGTGGTGGATAGTGAACCTAGACCCA 2046
Db 6983 TCGCCTCCAGAACGTGGCTGTGGTGGTGGATAGTGAACCTAGACCCA 7042
Qy 2047 ACAGAGTGAATGTGGATGCTCTCTGATCGATGAAGATGGAACCTGACCTGCA 2106
Db 7043 ACAGAGTGAATGTGGATGCTCTCTGATCGATGAAGATGGAACCTGACCTGCA 7102
Qy 2107 AGGTGGGCTATGTGGAGATGGCTTCTCATGAGTGGAACTGTGTCAGGCTCCTGATGT 2166
Db 7103 AGGTGGGCTATGTGGAGATGGCTTCTCATGAGTGGAACTGTGTCAGGCTCCTGATGT 7162
Qy 2167 CTTCCCTCTACTCAAACTTCTGACGGAAGTGTGGCTATTTCCAAACAGCTCAGCTC 2226
Db 7163 CTTCCCTCTACTCAAACTTCTGACGGAAGTGTGGCTATTTCCAAACAGCTCAGCTC 7222
Qy 2227 GAGGCGGTGCAATTTAGAACACCTGACTGACCTGTCCATCGCGGACCCCTTTGTGC 2286
Db 7223 GAGGCGGTGCAATTTAGAACACCTGACTGACCTGTCCATCGCGGACCCCTTTGTGC 7282
Qy 2287 CACAAACAGTGGGCTGGGGAGATGAGACCTGTCTGGGGGAGACATCGAGACCAACC 2346
Db 7283 CACAAACAGTGGGCTGGGGAGATGAGACCTGTCTGGGGGAGACATCGAGACCAACC 7342
Qy 2347 TCGCAATGTGAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACTCGCAACGA 2406
Db 7343 TCGCAATGTGAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACTCGCAACGA 7402
Qy 2407 GGTTGGGAGCAAGTGTCTATCTGACGAGGACCCATCAACCGAGGAGACA 2466
Db 7403 GGTTGGGAGCAAGTGTCTATCTGACGAGGACCCATCAACCGAGGAGACA 7462
Qy 2467 GGTGTTGTTGATGGAAGAGCAATCTGACGTGGGACATCTTTGCCCTCCCAATGGGATCTTC 2526
Db 7463 GGTGTTGTTGATGGAAGAGCAATCTGACGTGGGACATCTTTGCCCTCCCAATGGGATCTTC 7522
Qy 2527 ATGTCATTTCCAGGCTTTAAAGACACCCCTGCCCGGTCACCTTGACCCACACTGGCT 2586
Db 7523 ATGTCATTTCCAGGCTTTAAAGACACCCCTGCCCGGTCACCTTGACCCACACTGGCT 7582
Qy 2587 TGGGAGCAGGATCTCTTTGCCATCATCTGCTGACCTGGGCTGTGCTTGGCTGCTT 2646

Db 7583 TGGGAGCAGGATCTTTTGGCCATCATCTTGGTGACCTGGGGCTGTGCTTGGCTGCTT 7642
Qy 2647 ACTCTACTTTCCGATAAACCAGAGAACAAATCGGCTTCCAGCATTTTTCAGTCGGAAGAGG 2706
Db 7643 ACTCTACTTTCCGATAAACCAGAGAACAAATCGGCTTCCAGCATTTTTCAGTCGGAAGAGG 2702
Qy 2707 ACATTAAATGTTGAGCTTCTGGCAAGCAGAGCCTGAGAAATATCTCGAAACCCCTTTGATG 2766
Db 7703 ACATTAAATGTTGAGCTTCTGGCAAGCAGAGCCTGAGAAATATCTCGAAACCCCTTTGATG 2762
Qy 2767 AGAGCACAACCTCAGCTCCCCAGAACCTTCTAGACCCCTTCACGACCTCGAAGAAC 2826
Db 7763 AGAGCACAACCTCAGCTCCCCAGAACCTTCTAGACCCCTTCACGACCTCGAAGAAC 7822
Qy 2827 GGCAGCTTGGGGCAATGACCCCTTGGAGCACTGTGAGGGCCTGGACGGGAGATGCCAG 2886
Db 7823 GGCAGCTTGGGGCAATGACCCCTTGGAGCACTGTGAGGGCCTGGACGGGAGATGCCAG 7882
Qy 2887 CCATCATCTACCTGCGACCTGGGCCATCACTGTGAATTTCTCAGCACAGTTGCTTTTATG 2946
Db 7883 CCATCATCTACCTGCGACCTGGGCCATCACTGTGAATTTCTCAGCACAGTTGCTTTTATG 7942
Qy 2947 GAACGTAAAGTCTTTAAGCACTCAGAGCACTCAGAGCCATACCTCTCTGCTGATCTGGGG 3006
Db 7943 GAACGTAAAGTCTTTAAGCACTCAGAGCACTCAGAGCCATACCTCTCTGCTGATCTGGGG 8002
Qy 3007 TTGTTTCTGTGGTGAGAGATGTGCTGTGCGCCACCCAGTACAGCTTCTCTCTCTGAC 3066
Db 8003 TTGTTTCTGTGGTGAGAGATGTGCTGTGCGCCACCCAGTACAGCTTCTCTCTCTGAC 8062
Qy 3067 CTTTGGCTCTTCTCTTGTACTCTTCACTGCGACCTGCTCATTCTGCGCTTACATG 3126
Db 8063 CTTTGGCTCTTCTCTTGTACTCTTCACTGCGACCTGCTCATTCTGCGCTTACATG 8122
Qy 3127 ATGGTAACTGTGATCTTCTTCCCTGTAGATGTAAGCTCCTTTGTATCCCAAGC 3186
Db 8123 ATGGTAACTGTGATCTTCTTCCCTGTAGATGTAAGCTCCTTTGTATCCCAAGC 8182
Qy 3187 CCCTAGCCAGTCTGTGACACAGGAACCTGTGCAAAATAAGGTTTATGGACAGAAAAA 3246
Db 8183 CCCTAGCCAGTCTGTGACACAGGAACCTGTGCAAAATAAGGTTTATGGACAGAAAAA 8242
Qy 3247 AAAAAAAA 3254
Db 8243 AGTCAACA 8250

RESULT 5
AK074051LOCUS Homo sapiens mRNA for FLJ00122 protein.
DEFINITION

ACCESSION AK074051

VERSION AK074051.1 GI:18676449

KEYWORDS fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.

The nucleotide sequence of a long cDNA clone isolated from human

spleen

Published Only in Database (2002)

2 (bases 1 to 5604)

Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

JOURNAL

COMMENT

NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert and 5'- &

3'-end one pass sequencing: Research Association for Biotechnology;
cDNA library construction and clone selection: Kazusa DNA Research
Institute.

FEATURES

Location/Qualifiers

1..5604
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FLJ00122"
/tissue_type="spleen"
/dev_stage="adult"
/note="vector:pBluescriptII SK plus"
1..5604
/gene="FLJ00122"
41..5213
/gene="FLJ00122"
/note="Start codon is not identified."
/codon_start=3
/evidence="not experimental"
/product="FLJ00122 protein"
/protein_id="BAB84877.1"
/db_xref="GI:18676450"
/translation="DGSAGRLCDKQTSACGPVQVQCHIHATCEYNGNTASCTCKAYE
GDTLCSEMDPCTGLTPGGCSRNAEIKTGTHTCVCOQWGTNGRDCSEINNCLLP
SAGGCHDNASCLYVPGQNECEKKGFRNGIDCEPITSCLEQTKCHPLASQCSST
GWSVCVCGDYGDFLCYGNAAVELSFLSEAAIENRWNNASLOPITLSATSNLTIV
PSQATEMDQDEKSPWLSQSNIPALIKYHMLLGYRVADLOTLSASDMLATSLQNP
LHLAKVDGNITIGASIVDGNAAINGVIHINKVLVPPORLUTSLPMLLMLQMPD
YSIFGYIIOYLNAAIEADAYTFAPNNNAIENYIEKKVLSLEEDVLRVHVLEE
KLLKNDLHNGHREMLTFGSYFLFFLHNDQIYVNEAPINYTNVATDKGVIHGLKVL
ETQKRCNNDDTIIRGCRCTCSBLTCTPGFKTSLGNEKRCIYTSYFMRRTLPFG
OPKVRTVITRECCAGFGPQPCPGNAONVCFNGICLDGVNGTGVCECEGFSGT
ACETCEKGJHCDQACSVHRCNOCPGLGSDCDGVHGVCDNATTEDCNGT
CHTSANLNSGTSCAKCAAGFQNGTICTAINACEISNGGCSAKCKRTIPGRV
CTCKAGYTDGIVCLIEPNCLEHGCDKNAECTQTGPNQAAACNLPAVTDGKVTLL
INVLCTKNGGSEFAICNHTGQVVERTCTCPNYIYGDFTCRGSIYQELPKNPKTSQYF
FQLOHFHVLDVLPQGFVFAFLSAADFDEARVQDWKYGKLPQVLRVHVACHQLL
ENKLSINATSLQGLTIVISVQSIVYINNKATISSDIISTNGIVHIIDKLLSPKNI
LITPKDNRILONLITLATNGNIVKFSNLIODSGLLSVITDPIHTPVLFPWTDQAL
HALPAEQDFLEQNQDKOKLEYLKFHVIRDAKVLAVDLPTSTAKWTLQGSLSVKCG
ACNDTPSPRNSKPGKQKOLYNLPFRNLEGRERCSLVIOIPRCCKGYFGRDQCA
CPGDPDAPNCVGLCDQYSATCECKNTGNTGNTACMCWPGRFGPDCPLPCGCSDBHG
CDDGITGSQCLCTGTGTGTCSTQAVLPACTPCSAHATCKENNTCBNDLDEYDG
ITCTVDFCKQDNGCAKVARCSOKGTVKSCOKYKGDGHSCTEIDPCADLNGGC
HEHATCKTPGKHCKECSHYVGLNCEPEQLPIDRCLQDNGQCHADAKCVDLHFQ
DITVGVHLRSLPQKLTDFKAREACANEATMATYNQLSIAQAKYHLCSAGWLET
GRVAYTPAFASQNGSGVGVYDYGPRNPKSEMWDVFCYRMDKVNCTCKVYGVGDFS
CSGNLQVLMSPFSLTNFLTEVLYSNSARGAFLEHLTDLISIRGTLFVPQNSGLGE
NETLSGRDI EHLANVSMFPYNDLVNGTTLQTRIGSKLITASQDLPQLOPTEFRVDFG
ATLQWDIFASNGI IHVISRPLKAPAPVTLTHTGLGAGIFFALILVTGAVALAAYSYF
RINRRTIGFQHPSESEDINVAALGKQKQPENISNPLYESTTSAPRPSYDPFDTDSBERQ
LSENDPLRLT"

gene

CDS

Db	Qy	2536	2537	2538	2539	2540	2541	2542	2543	2544	2545	2546	2547	2548	2549	2550	2551	2552	2553	2554	2555	2556	2557	2558	2559	2560	2561	2562	2563	2564	2565	2566	2567	2568	2569	2570	2571	2572	2573	2574	2575	2576	2577	2578	2579	2580	2581	2582	2583	2584	2585	2586	2587	2588	2589	2590	2591	2592	2593	2594	2595	2596	2597	2598	2599	2600	2601	2602	2603	2604	2605	2606	2607	2608	2609	2610	2611	2612	2613	2614	2615	2616	2617	2618	2619	2620	2621	2622	2623	2624	2625	2626	2627	2628	2629	2630	2631	2632	2633	2634	2635	2636	2637	2638	2639	2640	2641	2642	2643	2644	2645	2646	2647	2648	2649	2650	2651	2652	2653	2654	2655	2656	2657	2658	2659	2660	2661	2662	2663	2664	2665	2666	2667	2668	2669	2670	2671	2672	2673	2674	2675	2676	2677	2678	2679	2680	2681	2682	2683	2684	2685	2686	2687	2688	2689	2690	2691	2692	2693	2694	2695	2696	2697	2698	2699	2700	2701	2702	2703	2704	2705	2706	2707	2708	2709	2710	2711	2712	2713	2714	2715	2716	2717	2718	2719	2720	2721	2722	2723	2724	2725	2726	2727	2728	2729	2730	2731	2732	2733	2734	2735	2736	2737	2738	2739	2740	2741	2742	2743	2744	2745	2746	2747	2748	2749	2750	2751	2752	2753	2754	2755	2756	2757	2758	2759	2760	2761	2762	2763	2764	2765	2766	2767	2768	2769	2770	2771	2772	2773	2774	2775	2776	2777	2778	2779	2780	2781	2782	2783	2784	2785	2786	2787	2788	2789	2790	2791	2792	2793	2794	2795	2796	2797	2798	2799	2800	2801	2802	2803	2804	2805	2806	2807	2808	2809	2810	2811	2812	2813	2814	2815	2816	2817	2818	2819	2820	2821	2822	2823	2824	2825	2826	2827	2828	2829	2830	2831	2832	2833	2834	2835	2836	2837	2838	2839	2840	2841	2842	2843	2844	2845	2846	2847	2848	2849	2850	2851	2852	2853	2854	2855	2856	2857	2858	2859	2860	2861	2862	2863	2864	2865	2866	2867	2868	2869	2870	2871	2872	2873	2874	2875	2876	2877	2878	2879	2880	2881	2882	2883	2884	2885	2886	2887	2888	2889	2890	2891	2892	2893	2894	2895	2896	2897	2898	2899	2900	2901	2902	2903	2904	2905	2906	2907	2908	2909	2910	2911	2912	2913	2914	2915	2916	2917	2918	2919	2920	2921	2922	2923	2924	2925	2926	2927	2928	2929	2930	2931	2932	2933	2934	2935	2936	2937	2938	2939	2940	2941	2942	2943	2944	2945	2946	2947	2948	2949	2950	2951	2952	2953	2954	2955	2956	2957	2958	2959	2960	2961	2962	2963	2964	2965	2966	2967	2968	2969	2970	2971	2972	2973	2974	2975	2976	2977	2978	2979	2980	2981	2982	2983	2984	2985	2986	2987	2988	2989	2990	2991	2992	2993	2994	2995	2996	2997	2998	2999	3000	3001	3002	3003	3004	3005	3006	3007	3008	3009	3010	3011	3012	3013	3014	3015	3016	3017	3018	3019	3020	3021	3022	3023	3024	3025	3026	3027	3028	3029	3030	3031	3032	3033	3034	3035	3036	3037	3038	3039	3040	3041	3042	3043	3044	3045	3046	3047	3048	3049	3050	3051	3052	3053	3054	3055	3056	3057	3058	3059	3060	3061	3062	3063	3064	3065	3066	3067	3068	3069	3070	3071	3072	3073	3074	3075	3076	3077	3078	3079	3080	3081	3082	3083	3084	3085	3086	3087	3088	3089	3090	3091	3092	3093	3094	3095	3096	3097	3098	3099	3100	3101	3102	3103	3104	3105	3106	3107	3108	3109	3110	3111	3112	3113	3114	3115	3116	3117	3118	3119	3120	3121	3122	3123	3124	3125	3126	3127	3128	3129	3130	3131	3132	3133	3134	3135	3136	3137	3138	3139	3140	3141	3142	3143	3144	3145	3146	3147	3148	3149	3150	3151	3152	3153	3154	3155	3156	3157	3158	3159	3160	3161	3162	3163	3164	3165	3166	3167	3168	3169	3170	3171	3172	3173	3174	3175	3176	3177	3178	3179	3180	3181	3182	3183	3184	3185	3186	3187	3188	3189	3190	3191	3192	3193	3194	3195	3196	3197	3198	3199	3200	3201	3202	3203	3204	3205	3206	3207	3208	3209	3210	3211	3212	3213	3214	3215	3216	3217	3218	3219	3220	3221	3222	3223	3224	3225	3226	3227	3228	3229	3230	3231	3232	3233	3234	3235	3236	3237	3238	3239	3240	3241	3242	3243	3244	3245	3246	3247	3248	3249	3250	3251	3252	3253	3254	3255	3256	3257	3258	3259	3260	3261	3262	3263	3264	3265	3266	3267	3268	3269	3270	3271	3272	3273	3274	3275	3276	3277	3278	3279	3280	3281	3282	3283	3284	3285	3286	3287	3288	3289	3290	3291	3292	3293	3294	3295	3296	3297	3298	3299	3300	3301	3302	3303	3304	3305	3306	3307	3308	3309	3310	3311	3312	3313	3314	3315	3316	3317	3318	3319	3320	3321	3322	3323	3324	3325	3326	3327	3328	3329	3330	3331	3332	3333	3334	3335	3336	3337	3338	3339	3340	3341	3342	3343	3344	3345	3346	3347	3348	3349	3350	3351	3352	3353	3354	3355	3356	3357	3358	3359	3360	3361	3362	3363	3364	3365	3366	3367	3368	3369	3370	3371	3372	3373	3374	3375	3376	3377	3378	3379	3380	3381	3382	3383	3384	3385	3386	3387	3388	3389	3390	3391	3392	3393	3394	3395	3396	3397	3398	3399	3400	3401	3402	3403	3404	3405	3406	3407	3408	3409	3410	3411	3412	3413	3414	3415	3416	3417	3418	3419	3420	3421	3422	3423	3424	3425	3426	3427	3428	3429	3430	3431	3432	3433	3434	3435	3436	3437	3438	3439	3440	3441	3442	3443	3444	3445	3446	3447	3448	3449	3450	3451	3452	3453	3454	3455	3456	3457	3458	3459	3460	3461	3462	3463	3464	3465	3466	3467	3468	3469	3470	3471	3472	3473	3474	3475	3476	3477	3478	3479	3480	3481	3482	3483	3484	3485	3486	3487	3488	3489	3490	3491	3492	3493	3494	3495	3496	3497	3498	3499	3500	3501	3502	3503	3504	3505	3506	3507	3508	3509	3510	3511	3512	3513	3514	3515	3516	3517	3518	3519	3520	3521	3522	3523	3524	3525	3526	3527	3528	3529	3530	3531	3532	3533	3534	3535	3536	3537	3538	3539	3540	3541	3542	3543	3544	3545	3546	3547	3548	3549	3550	3551	3552	3553	3554	3555	3556	3557	3558	3559	3560	3561	3562	3563	3564	3565	3566	3567	3568	3569	3570	3571	3572	3573	3574	3575	3576	3577	3578	3579	3580	3581	3582	3583	3584	3585	3586	3587	3588	3589	3590	3591	3592	3593	3594	3595	3596	3597	3598	3599	3600	3601	3602	3603	3604	3605	3606	3607	3608	3609	3610	3611	3612	3613	3614	3615	3616	3617	3618	3619	3620	3621	3622	3623	3624	3625	3626	3627	3628	3629	3630	3631	3632	3633	3634	3635	3636	3637	3638	3639	3640	3641	3642	3643	3644	3645	3646	3647	3648	3649	3650	3651	3652	3653	3654	3655	3656	3657	3658	3659	3660	3661	3662	3663	3664	3665	3666	3667	3668	3669	3670	3671	3672	3673	3674	3675	3676	3677	3678	3679	3680	3681	3682	3683	3684	3685	3686	3687	3688	3689	3690	3691	3692	3693	3694	3695	3696	3697	3698	3699	3700	3701	3702	3703	3704	3705	3706	3707	3708	3709	3710	3711	3712	3713	3714	3715	3716	3717	3718	3719	3720	3721	3722	3723	3724	3725	3726	3727	3728	3729	3730	3731	3732	3733	
----	----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--

QY 1327 CQTGTGACACTCAGGAGTTTTCCTGCAAGTGTGACGCTCCTTTGTTCTGCTCATGCCA 1386
DB |||||
QY 3676 CQTGTGACACTCAGGAGTTTTCCTGCAAGTGTGACGCTCCTTTGTTCTGCTCATGCCA 3735
DB |||||
QY 1387 CQTGTGAGAGAACACACGCTGAGTGTAACTGGATTTATCAAGTGTGACGAATCACAT 1446
DB |||||
QY 3736 CQTGTGAGAGAACACACGCTGAGTGTAACTGGATTTATCAAGTGTGACGAATCACAT 3795
DB |||||
QY 1447 GCACAGTTGTGGATTTCTGCAAAACAGGACAAACGCGGGCTGTGCAAAAGTGGCCAGATGCT 1506
DB |||||
QY 3796 GCACAGTTGTGGATTTCTGCAAAACAGGACAAACGCGGGCTGTGCAAAAGTGGCCAGATGCT 3855
DB |||||
QY 1507 CCCAGAGGACAGAAAGTCTCTCTGAGCTGCCAGAGGATACAAAGGGGACGCGGCACA 1566
DB |||||
QY 3856 CCCAGAGGACAGAAAGTCTCTCTGAGCTGCCAGAGGATACAAAGGGGACGCGGCACA 3915
DB |||||
QY 1567 GCTGCACAGAGATAGACCCCTGTGACAGCGGCTTTAAACGAGGGTGTACAGACAGCGCA 1626
DB |||||
QY 3916 GCTGCACAGAGATAGACCCCTGTGACAGCGGCTTTAAACGAGGGTGTACAGACAGCGCA 3975
DB |||||
QY 1627 CQTGTGAGATGACAGCCCGGGCAAGCAACAGTGTGAGTGTAAAGTCACTATGTCCGAG 1686
DB |||||
QY 3976 CQTGTGAGATGACAGCCCGGGCAAGCAACAGTGTGAGTGTAAAGTCACTATGTCCGAG 4035
DB |||||
QY 1687 ATGGGCTGAACGTGTGAGCGGAGAGAGTGCCTCATTTGACCGCTGTACAGCAATGGGC 1746
DB |||||
QY 4036 ATGGGCTGAACGTGTGAGCGGAGAGAGTGCCTCATTTGACCGCTGTACAGCAATGGGC 4095
DB |||||
QY 1747 AGTGCCATGACAGACGCCAAATGTGTGACGCTTCACTTCCAGATACCACTGTGGGGTGT 1806
DB |||||
QY 4096 AGTGCCATGACAGACGCCAAATGTGTGACGCTTCACTTCCAGATACCACTGTGGGGTGT 4155
DB |||||
QY 1807 TCCATCTACGCTCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCT 1866
DB |||||
QY 4156 TCCATCTACGCTCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCT 4215
DB |||||
QY 1867 GTGCCAAGAGCTGCGACATGCGCAACCTACAAACAGCTCTCCATGCCAGAGGCGCA 1926
DB |||||
QY 4216 GTGCCAAGAGCTGCGACATGCGCAACCTACAAACAGCTCTCCATGCCAGAGGCGCA 4275
DB |||||
QY 1927 AGTACCACCTGTGCTCAGCAGCTGGCTGGAGACCGGGGGTGCCTTACCCACAGGCT 1986
DB |||||
QY 4276 AGTACCACCTGTGCTCAGCAGCTGGCTGGAGACCGGGGGTGCCTTACCCACAGGCT 4335
DB |||||
QY 1987 TCGCTCCCAAGACTGTGGCTCTGGTGTGGTGGATAGTGGACTATGGAACCTAGACCCA 2046
DB |||||
QY 4336 TCGCTCCCAAGACTGTGGCTCTGGTGTGGTGGATAGTGGACTATGGAACCTAGACCCA 4395
DB |||||
QY 2047 ACAAGAGTGAATGTGGATGTCTTCTGCTATCGGATGAAGATGTGAACCTGCACTGCA 2106
DB |||||
QY 4396 ACAAGAGTGAATGTGGATGTCTTCTGCTATCGGATGAAGATGTGAACCTGCACTGCA 4455
DB |||||
QY 2107 AGTGGGGCTATGTGGAGATGGCTTCTCATGCGAGTGGAACTGCTGCGAGCTCTGATGT 2166
DB |||||
QY 4456 AGTGGGGCTATGTGGAGATGGCTTCTCATGCGAGTGGAACTGCTGCGAGCTCTGATGT 4515
DB |||||
QY 2167 CCTTCCCTCACTCACAAACTTCTGACGGAAGTGTGCTGGCTATTTCCAAACAGCTCAGCTC 2226
DB |||||
QY 4516 CCTTCCCTCACTCACAAACTTCTGACGGAAGTGTGCTGGCTATTTCCAAACAGCTCAGCTC 4575
DB |||||
QY 2227 GAGGCGGTGCAATTTCTAGAACACTGACTGCTGCTCATCGCGGACACCTCTTTGTGC 2286
DB |||||
QY 4576 GAGGCGGTGCAATTTCTAGAACACTGACTGCTGCTCATCGCGGACACCTCTTTGTGC 4635
DB |||||
QY 2287 CACAGACAGTGGGCTGGGGAGATGAGACTTGTCTGGGGGGGACATCGACACACACC 2346
DB |||||
QY 4636 CACAGACAGTGGGCTGGGGAGATGAGACTTGTCTGGGGGGGACATCGACACACACC 4695
DB |||||
QY 2347 TCGCCAAATGTGAGCATGTTTTTCTACATGACCTTGTCAATGGGACCAACCTCTGCAAGCA 2406
DB |||||
QY 4696 TCGCCAAATGTGAGCATGTTTTTCTACATGACCTTGTCAATGGGACCAACCTCTGCAAGCA 4755
DB |||||

QY 2407 GGGTGGAGAGCAAGCTGTCTCATCTGCGAGCAGGACCCACTCCAAACCGACGAGACCA 2466
DB |||||
QY 4756 GGGTGGAGAGCAAGCTGTCTCATCTGCGAGCAGGACCCACTCCAAACCGACGAGACCA 4815
DB |||||
QY 2467 GGGTGGTGTGAGAGAGCAATCTGCAAGTGGGACATCTTTGCTCCCAATGGGATCATTC 2526
DB |||||
QY 4816 GGGTGGTGTGAGAGAGCAATCTGCAAGTGGGACATCTTTGCTCCCAATGGGATCATTC 4875
DB |||||
QY 2527 ATGTCAATTTCCAGGCTTTTAAAGACACCCCTGCCCGCTGACCTTGACCACTGGCT 2586
DB |||||
QY 4876 ATGTCAATTTCCAGGCTTTTAAAGACACCCCTGCCCGCTGACCTTGACCACTGGCT 4935
DB |||||
QY 2587 TGGGAGCAGGAGATCTTTTGGCCATCATCTGCTGACTGGGGCTGTGCTTTGGCTGCTT 2646
DB |||||
QY 4936 TGGGAGCAGGAGATCTTTTGGCCATCATCTGCTGACTGGGGCTGTGCTTTGGCTGCTT 4995
DB |||||
QY 2647 ACTCTACTCTTCGGATTAACCGGAGAACATCGGCTTCAGCAATTTGAGTCGGAAGAG 2706
DB |||||
QY 4996 ACTCTACTCTTCGGATTAACCGGAGAACATCGGCTTCAGCAATTTGAGTCGGAAGAG 5055
DB |||||
QY 2707 ACATTAATGTTGAGCTCTTGGCAAGCAGCAGCTGAGAATATCTCGAACCCCTTGATG 2766
DB |||||
QY 5056 ACATTAATGTTGAGCTCTTGGCAAGCAGCAGCTGAGAATATCTCGAACCCCTTGATG 5115
DB |||||
QY 2767 AGAGCACAACTCTAGCTCTCCCGCAGAACCTTCTTACGACCCCTTCAAGGACTCTGAAGAC 2826
DB |||||
QY 5116 AGAGCACAACTCTAGCTCTCCCGCAGAACCTTCTTACGACCCCTTCAAGGACTCTGAAGAC 5175
DB |||||
QY 2827 GGCAGCTTGAGGGGCAATGACCCCTTGAGGACATCTGTGAGGGGCTGGACGGGAGATGCCAG 2886
DB |||||
QY 5176 GGCAGCTTGAGGGGCAATGACCCCTTGAGGACATCTGTGAGGGGCTGGACGGGAGATGCCAG 5235
DB |||||
QY 2887 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTTCTCAGCACCAAGTTGCCCTTTAG 2946
DB |||||
QY 5236 CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTTCTCAGCACCAAGTTGCCCTTTAG 5295
DB |||||
QY 2947 GAACGTAAGTCTTTTAAAGCACTCAGAACCCATACCTCATCTCTGCTGATCTGCGGG 3006
DB |||||
QY 5296 GAACGTAAGTCTTTTAAAGCACTCAGAACCCATACCTCATCTCTGCTGATCTGCGGG 5355
DB |||||
QY 3007 TTGTTCTGTGGGTGAGAGATGTGTGCTGCGCCACCCAGTACAGCTTCTCTCTGAC 3066
DB |||||
QY 5356 TTGTTCTGTGGGTGAGAGATGTGTGCTGCGCCACCCAGTACAGCTTCTCTCTGAC 5415
DB |||||
QY 3067 CCTTTGGCTCTTCTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3126
DB |||||
QY 5416 CCTTTGGCTCTTCTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5475
DB |||||
QY 3127 ATGGGTAACCTGTGATCTTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3186
DB |||||
QY 5476 ATGGGTAACCTGTGATCTTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5535
DB |||||
QY 3187 CCCTAGCCCACTGCTGACACAGGAACTGTGCACAATAAGGTTTATGGAAACAGAAAAA 3246
DB |||||
QY 5536 CCCTAGCCCACTGCTGACACAGGAACTGTGCACAATAAGGTTTATGGAAACAGAAAAA 5595
DB |||||
QY 3247 AAAAAAAAA 3254
DB |||||
QY 5596 AGTCAACA 5603
DB |||||

RESULT 6

AK024503
LOCUS AK024503 4575 bp mRNA linear PRI 29-SEP-2000
DEFINITION Homo sapiens mRNA for FLJ00112 protein, partial cds.
ACCESSION AK024503
VERSION AK024503.1 GI:10440525
KEYWORDS fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4575)

AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
 TITLE The nucleotide sequence of a long cDNA clone isolated from human spleen
 JOURNAL Published Only in DataBase (2000)
 REFERENCE 2 (bases 1 to 4575)
 AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
 TITLE Direct Submission
 JOURNAL Submitted (24-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan [E-mail:cdmainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913, Fax:81-438-52-3914]
 COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- and 3'-end one pass sequencing and clone selection: Kazusa DNA Research Institute.
 FEATURES
 Location/Qualifiers
 1..4575
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="as00112"
 /tissue_type="spleen"
 /dev_stage="adult"
 /note="vector:pBluescriptII SK plus"
 gene
 1..4575
 /gene="FLJ00112"
 CDS
 1..4184
 /note="Start codon is not identified."
 /codon_start=1
 /product="FLJ00112 protein"
 /protein_id="BAB15793.1"
 /db_xref="GI:10440526"
 /translation="WHILFGMSDGTGVCCEGEGSGTACETCTEGKYGIHCDQACSVH GRNCGPLDGGDCDVGWRGVHCDNATDNCNGTCHTSANLNSDGTASCKCAAG POGNGTICTAINACELISNGCSAKADCKETTPBRVCTCKAGYTGDIIVCLEINPCLE NHGCDCKNAECTGTGNQACNCLPAYDGTGKVTCLINVLKPNKGCCSFAICNHTGP VERTCTKACNYIGDGTGSGIYQELFPKPTQSYFFQJQEHFVKDLVDFGPFVTFAP LSAADFEARVDKMDKYLMPQVLRHVHACHQLLENKLI.SNATSLQGEPIVISVS QSTVYNNKAKIISDDIISTINGIIVHIDKLLSPKNLLITPKDNGSRILQNLTLATNN GYIKFHLVRADKVLAVDLPTSTAKTLQGSSELVSKCAGARDIDGLFNGOTCRIVREL LFDLGVAYGIDCLLDIPTLGGRCDFTEFDASGECSGVNTPSCPNWSKPKGVOKCL YNLPFRNLEGRCESLVIQIPRCNGIFGRDCQACPGDPAPNNRNVGLDQISAT GBCKNTGNGTACEMCFWRFPGDCLPCGSDHGDGDIITGSGQCLCTETGWTGSC DTQAVLPVCTPPCSAHATCKENNTCECNLYEGDGIITCTVDFCKQDNGGCAKVARC SOKGTVCSCSKYKGDGHSCTEIDPCADLNGGCHHEATCTVDFCKQDNGGCAKVARC VGDGLNCEPOLPIDRCLDNGOCHADAKVDLHFDOTTVGVFHLRSPLGOVKLTPDK AREACANEATMYNQLSYAQKAKYHLCAGWLETGRVAYPTAFASONGSGSVLIEV DYGPENKSEMDFVCTRMKDVNCTCKVGTGVDGFSGSLNQLQVLMSPFSLNLFTEV LAYSNSARGRAFLLEHLTDLISLNGTLFVFNQSLGENETLSGRDIBHHLANVSMFPYN DIVNGTTLQRLGSKLLITASDPLQFETRFVDGRAILQWDIFASNGIIVHSRLK APPAPVTLHTGLAGIFPAILVTVGAVALAAYSYPINRRTIGFQHFSEEDINVAALGKQQFENISNPLYESTTSAPPEPSYDPTDSEERQLEGNDPLRTL"

ORIGIN

Query Match 99.1%; Score 3231; DB 9; Length 4575;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 3237; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 7 AGCAGGAGCTTCCCAAGAACCCGAAACTTCCAGATTTCTTCCAGTTGACAGGACATT 66
 1327 ATCAGGAGCTTCCCAAGAACCCGAAACTTCCAGATTTCTTCCAGTTGACAGGACCT 1386
 67 TCGTGAAGATCTGGTGGCGCCAGGCCCTTCACTGTTTGTGACCTTTATCTGCAGCCT 126
 1387 TCGTGAAGATCTGGTGGCGCCAGGCCCTTCACTGTTTGTGACCTTTATCTGCAGCCT 1446
 127 TTGATGAGGAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCGGCTTCTTC 186
 1447 TTGATGAGGAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCGGCTTCTTC 1506

187 GGTACCATGTGTCGCTGCGCCACACGAGCTGCTTCTTGAAAAACCTGAAATTTGATCTCAAAATG 246
 1507 GGTACCATGTGTCGCTGCGCCACACGAGCTGCTTCTTGAAAAACCTGAAATTTGATCTCAAAATG 1566
 247 CTACTTCCCTCCAGAGGAGAGCCAAATAGTCACTCCGCTCTCTCAGAGCAGCGGTGATATAA 306
 1567 CTACTTCCCTCCAGAGGAGAGCCAAATAGTCACTCCGCTCTCTCAGAGCAGCGGTGATATAA 1626
 307 ATAAATAGCGCTAAGATCATATCCAGTGATATCATCAGTACTATCAGTATGGGATGTTGATATCA 366
 1627 ACAATTAAGGCTAAGATCATATCCAGTGATATCATCAGTACTATCAGTATGGGATGTTGATATCA 1686
 367 TAGACAAATTTGCTATCTCCAAAAATTTGCTTATCATCTCCCAAGACAACTCTGGAAGAA 426
 1687 TAGACAAATTTGCTATCTCCAAAAATTTGCTTATCATCTCCCAAGACAACTCTGGAAGAA 1746
 427 TTCTGCAAAATCTTACGACTTTTGGCAACAAACAATGGCTACATCAATTTAGCAACTTAA 486
 1747 TTCTGCAAAATCTTACGACTTTTGGCAACAAACAATGGCTACATCAATTTAGCAACTTAA 1806
 487 TACAGGACTCAGGTTTGTGAGTGTATCACCAGATCCCATCCACACCCAGTCACTCTCT 546
 1807 TACAGGACTCAGGTTTGTGAGTGTATCACCAGATCCCATCCACACCCAGTCACTCTCT 1866
 547 TCTGGCCCAACGACCAAGCCCTCCATGCCCTACCTGCTGAAACAACAGGACTTCTCTGTTCA 606
 1867 TCTGGCCCAACGACCAAGCCCTCCATGCCCTACCTGCTGAAACAACAGGACTTCTCTGTTCA 1926
 607 ACCAAGACAAACAGGACAAAGCTGAAAGGATATTTGAAGTTTCAATGTGATACAGATGCCA 666
 1927 ACCAAGACAAACAGGACAAAGCTGAAAGGATATTTGAAGTTTCAATGTGATACAGATGCCA 1986
 667 AGGTTTACGCTGTGATCTTCCACATCCACTGCCCTGGAAGACCTTCAAGGTTTCAGAGC 726
 1987 AGGTTTACGCTGTGATCTTCCACATCCACTGCCCTGGAAGACCTTCAAGGTTTCAGAGC 2046
 727 TGAGTGTGAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTTGAATGGCCAAACCT 786
 2047 TGAGTGTGAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTTGAATGGCCAAACCT 2106
 787 GCAGAAATGTGAGCGGGAGCTCTTGTGTAACCTGGGTGTGGCTTACGGCATTTGATGTC 846
 2107 GCAGAAATGTGAGCGGGAGCTCTTGTGTAACCTGGGTGTGGCTTACGGCATTTGATGTC 2166
 847 TCGTGAATTCATCCACCTGGGGGGCGCTGTGACACCTTTACTACTTTTCGATGCCTCGG 906
 2167 TCGTGAATTCATCCACCTGGGGGGCGCTGTGACACCTTTACTACTTTTCGATGCCTCGG 2226
 907 GGGAGTGTGGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG 966
 2227 GGGAGTGTGGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG 2286
 967 TGAAGCAGAAGTGTCTCTCAACCTGCCCTTCAAGAGAACCTTGAAGGCTGCGCGGAGC 1026
 2287 TGAAGCAGAAGTGTCTCTCAACCTGCCCTTCAAGAGAACCTTGAAGGCTGCGCGGAGC 2346
 1027 GGTGACGCTGTGTATACAGATCCCGAGGTGCTGCAAGGCTACTTTGGCGGAGACTGTC 1086
 2347 GGTGACGCTGTGTATACAGATCCCGAGGTGCTGCAAGGCTACTTTGGCGGAGACTGTC 2406
 1087 AGGCTTGGCTTGGAGGACAGATGCCCCGCTGTAAATACCGGGGTGTCTGCCTTGTAGT 1146
 2407 AGGCTTGGCTTGGAGGACAGATGCCCCGCTGTAAATACCGGGGTGTCTGCCTTGTAGT 2466
 1147 ACTCGGCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGAGCGGCTGTGAGATGT 1206
 2467 ACTCGGCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGAGCGGCTGTGAGATGT 2526
 1207 GCTGGCGGGGAGATTGGGGCTGATTTGTCTGCCCTGTGCTGCTCAGACCAACGACAGT 1266
 2527 GCTGGCGGGGAGATTGGGGCTGATTTGTCTGCCCTGTGCTGCTCAGACCAACGACAGT 2586

QY 1267 GCGATGATGGCATCAGGGCTCGGGCAGTGCCTCTGTGAACGGGGTGGACAGGCCCT 1326
DB |||||
QY 1327 CCGTGTGACACTCAGGCAGTCTTTCCTGCGAGTGTGTACGCTCCTTGTCTCTCATGCCA 1386
DB |||||
QY 1387 CCTGTAAGGAGAACACAGTGTGAGTGAACCTGTGATTTATGAAGGTGACGGAATCACAT 1446
DB |||||
QY 1447 GCACAGTGTGATTTCTGCAACAGGACAGCGGGGCTGTCCAAAGGTGGCCAGATGCT 1506
DB |||||
QY 1507 CCCAGAGGGCAGGAAGTCTCCTGAGCTGCGAGAGGATACAAAGGGGACGGGCACA 1566
DB |||||
QY 1567 GCTGACAGAGATAGACCCCTGTGACAGCGGCTTTAAGGAGGTGTCAAGAGCAGGCA 1626
DB |||||
QY 1627 CCTGTAAGATGACAGCGCGGCGGACAGCAGTGTGAGTGTAAAGTCACTATGTGCGAG 1686
DB |||||
QY 1687 ATGGGCTGAACTGTGAGCGGAGCAGCTGCCATTGACCGCTGTTCAGGACCAATGGGC 1746
DB |||||
QY 1747 AGTGCCATGACAGCGCCAAATGTGCGACCTCCACTTCCAGGATACCACTGTGCGGTGT 1806
DB |||||
QY 1807 TCCATCTAGCTCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGGCT 1866
DB |||||
QY 1867 GTGCCAACGAAGCTGCGACCACTGCAACCACTCTCTCTATGCCAGAGGCCA 1926
DB |||||
QY 1927 AGTACCACTGTGTCAGAGCTGGCTGGAGACCGGCGGCTTGCCTTACCCCAAGCCT 1986
DB |||||
QY 1987 TCGCTCCAGAACTGTGGGATGCTTCTGCTATCGGATGAAAGATGTGAACCTGCACTGCA 2046
DB |||||
QY 2047 ACAAGAGTGAATGTGGATGCTTCTGCTATCGGATGAAAGATGTGAACCTGCACTGCA 2106
DB |||||
QY 2107 AGTGGGTATGTGGAGATGGCTTCTCATGCAFGGGAACCTGTGAGTCTGTGATGT 2166
DB |||||
QY 2167 CCTTCCCTCACTCACAACCTTCTGACGGAAGTGTGGCTATTCGAAGCTCAGCTC 2226
DB |||||
QY 2227 GAGGCGGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGGGACCCCTCTTTGTGC 2286
DB |||||
QY 2287 CACAGACAGTGGCTGGGAGAGTATGACCTTGTCTGGCGGGACATCAGACCAACC 2346
DB |||||
QY 2347 TCGCCATGTGAGCATGTTTTTCTACATGACCTTGTCAATGGCAGCACCCCTGCAAGCA 2406
DB |||||

DB 3667 TCGCCAAATGTGAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACTTGCACCA 3726
QY 2407 GGGTGGGAGCAAGCTGCTCATCTACCTCCAGCAGGACCCACTCCAAACCGAGGACCA 2466
DB |||||
QY 2467 GGTGTTGTTGATGGAAGAGCCATTCGCACTGGGACATCTTTGCCCTCCAAATGGATCATTC 2526
DB |||||
QY 2527 ATGTCAATTTCCAGGCTTTTAAAGCACCCCTGCCCCCTGACCTTGACCCCACTGGCT 2586
DB |||||
QY 2587 TGGGAGCAGGATCTTCTTTCGCCATCATCTGCTGACCTGGGCTGTGGCTTGGCTGCTT 2646
DB |||||
QY 2647 ACTCCTACTTTCCGATATAACCGGAGAACAAATCGGCTTCAGCATTTTTCAGTCCGAAGAGG 2706
DB |||||
QY 2707 ACATTAATTTGCGAGCTCTTGGCAAGCAGCAGCTGAGAAATATCTCGAAACCCCTTGTATG 2766
DB |||||
QY 2767 AGAGCACAACTCTAGCTCCCGCAGAACCTTCTTACGACCCCTTACAGCACTCTGAGAAC 2826
DB |||||
QY 2827 GGCAGCTTTAGGCAATGACCCCTTTAGGACACTGTGAGGCGCTGCAAGGAGATGCCAG 2886
DB |||||
QY 2887 CCATCACTCACTGCCACCTGGGCGCATCAACTGTGAATTTCTCAGCACTGCTGCTTTTATG 2946
DB |||||
QY 2947 GAACTGAAGTCTTTTAAAGCTCAGAGCCTACCTCTCTGCTGCTGCTGCTGCTGCTGCTG 3006
DB |||||
QY 3007 TTGTTTCTGTGGGTGAGAGATGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3066
DB |||||
QY 3067 CCTTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 3126
DB |||||
QY 3127 ATGGCTAACTGTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 3186
DB |||||
QY 3187 CCCTAGCCAGTGGCTGACACAGGAACCTGTGCAACAATAAAGTTTATGGAACAGAAAAA 3246
DB |||||
QY 3247 AAAAAAAA 3254
DB 4567 AGTCAACA 4574

RESULT 7

AB052958

LOCUS

DEFINITION

AB052958

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

8214 bp mRNA linear PRI 11-SEP-2002

Homo sapiens mRNA for FELE-2, complete cds.

Homo sapiens

GI:22779440

Homo sapiens (human)

Homo sapiens

QY 967 TGAAGCAGAAGTGTCTCTACAACTGCGCCCTTCAAGAGGAACTGTGAAAGGCTGCGCGAGC 1026
Db |||||
QY 5941 TGAAGCAGAAGTGTCTCTACAACTGCGCCCTTCAAGAGGAACTGTGAAAGGCTGCGCGAGC 6000
Db |||||
QY 1027 GGTGAGCCTGTGTATACAGATCCCAAGTGTGTGCAAGGGTACTTTCGGGCGAGACTGTC 1086
Db |||||
QY 6001 GGTGAGCCTGTGTATACAGATCCCAAGTGTGTGCAAGGGTACTTTCGGGCGAGACTGTC 6060
Db |||||
QY 1087 AGGCTGCTCCCTGGAGGACAGATGCCCGTGTAAATAACCGGGGTGTCTGCTTGTATCAGT 1146
Db |||||
QY 6061 AGGCTGCTCCCTGGAGGACAGATGCCCGTGTAAATAACCGGGGTGTCTGCTTGTATCAGT 6120
Db |||||
QY 1147 ACTGGCCACCGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGGTGTGAGATGT 1206
Db |||||
QY 6121 ACTGGCCACCGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGGTGTGAGATGT 6180
Db |||||
QY 1207 GCTGGCGGGGAGATTCGGGGCTGATTCGTGCTGCTGTGGTGTCTCAGACACCGACAGT 1266
Db |||||
QY 6181 GCTGGCGGGGAGATTCGGGGCTGATTCGTGCTGCTGTGGTGTCTCAGACACCGACAGT 6240
Db |||||
QY 1267 GCGATGATGGCATCACCGGCTCCGGGCAAGTGCCTCTGTGAAACGGGGTGGACAGGCCCT 1326
Db |||||
QY 6241 GCGATGATGGCATCACCGGCTCCGGGCAAGTGCCTCTGTGAAACGGGGTGGACAGGCCCT 6300
Db |||||
QY 1327 CGTGTGACACTCAGGCAGTTCCTGCTGAGTGTGTACGCTCTCTGCTTGTCTGCTATGCGCA 1386
Db |||||
QY 6301 CGTGTGACACTCAGGCAGTTCCTGCTGAGTGTGTACGCTCTCTGCTTGTCTGCTATGCGCA 6360
Db |||||
QY 1387 CCTGTAAAGGAGAAACAACAGTGTGAGTGAACCTGGATTATGAAGGTGACCGGAATCACAT 1446
Db |||||
QY 6361 CCTGTAAAGGAGAAACAACAGTGTGAGTGAACCTGGATTATGAAGGTGACCGGAATCACAT 6420
Db |||||
QY 1447 GCAAGTGTGGATTTCTGAAAACAGGACAAACGGGGCTGTGCAAGGTGGCCAGATGCT 1506
Db |||||
QY 6421 GCAAGTGTGGATTTCTGAAAACAGGACAAACGGGGCTGTGCAAGGTGGCCAGATGCT 6480
Db |||||
QY 1507 CCCAGAGGACAGAGTCTCTGCGAGCTGCGAGAGGGATACAAAGGGACCGGACCA 1566
Db |||||
QY 6481 CCCAGAGGACAGAGTCTCTGCGAGCTGCGAGAGGGATACAAAGGGACCGGACCA 6540
Db |||||
QY 1567 GCTGCACAGATAGACCCCTGTGCGAGACGCCCTTAAACGGAGGGTGTCAACGACACGCCA 1626
Db |||||
QY 6541 GCTGCACAGATAGACCCCTGTGCGAGACGCCCTTAAACGGAGGGTGTCAACGACACGCCA 6600
Db |||||
QY 1627 CCTGTAAAGATACAGGCGCCCGGCAAGCAACAAGTGTGAGTGTAAAGTCACTATGTCGGAG 1686
Db |||||
QY 6601 CCTGTAAAGATACAGGCGCCCGGCAAGCAACAAGTGTGAGTGTAAAGTCACTATGTCGGAG 6660
Db |||||
QY 1687 ATGGGCTGAACTGTGAGCCGAGCAGCTGCCCATTTGACCGTGTCTTACAGACAAATGGGC 1746
Db |||||
QY 6661 ATGGGCTGAACTGTGAGCCGAGCAGCTGCCCATTTGACCGTGTCTTACAGACAAATGGGC 6720
Db |||||
QY 1747 AGTGCCATGACAGCGCCAAATGTGTGCACTTCCACTTCCAGGATACCACTGTGCGGGTGT 1806
Db |||||
QY 6721 AGTGCCATGACAGCGCCAAATGTGTGCACTTCCACTTCCAGGATACCACTGTGCGGGTGT 6780
Db |||||
QY 1807 TCCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCACAGAGGCT 1866
Db |||||
QY 6781 TCCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCACAGAGGCT 6840
Db |||||
QY 1867 GTGCCAAGAGCTGCGACCAATGCAACCTTACACAGCTCTCTATGCGCCAGAGGCCA 1926
Db |||||
QY 6841 GTGCCAAGAGCTGCGACCAATGCAACCTTACACAGCTCTCTATGCGCCAGAGGCCA 6900
Db |||||
QY 1927 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTCCTTACCCACAGCCT 1986
Db |||||
QY 6901 AGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTCCTTACCCACAGCCT 6960
Db |||||
QY 1987 TCGCCTCCAGAACTGTGGCTCTGGTGTGGTGTGGAGTAGTGGACTATGGACCTTAGACCCA 2046
Db |||||
QY 6961 TCGCCTCCAGAACTGTGGCTCTGGTGTGGTGTGGAGTAGTGGACTATGGACCTTAGACCCA 7020
Db |||||
QY 2047 ACAAGATGAATGTGGAGTGTCTTCTGCTATCGGATGAAGATGTGAATGCACTGCA 2106
Db |||||

Db |||||
QY 7021 ACAAGATGAATGTGGAGTGTCTTCTGCTATCGATGAAAGATGTGAACCTGCACTGCA 7080
Db |||||
QY 2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGAGTGGGAACTGTGTCAGGTCTGTATGT 2166
Db |||||
QY 7081 AGGTGGGCTATGTGGGAGATGGCTTCTCATGAGTGGGAACTGTGTCAGGTCTGTATGT 7140
Db |||||
QY 2167 CTTTCCCCTCACTCACAACTTCTGACGGAAGTGTGGCTATTTCCAAACAGCTCAGCTC 2226
Db |||||
QY 7141 CTTTCCCCTCACTCACAACTTCTGACGGAAGTGTGGCTATTTCCAAACAGCTCAGCTC 7200
Db |||||
QY 2227 GAGCGGTGCAATTTCTAGAACACCTGACTGCTGTCCATCCGGGCAACCTCTTTTGTGC 2286
Db |||||
QY 7201 GAGCGGTGCAATTTCTAGAACACCTGACTGACCTGTCCATCCGGGCAACCTCTTTTGTGC 7260
Db |||||
QY 2287 CACAGAACAGTGGCTGGGGAGATGAGACCTTGTCTGGGGGGGACATCGAGCACCA 2346
Db |||||
QY 7261 CACAGAACAGTGGCTGGGGAGATGAGACCTTGTCTGGGGGGGACATCGAGCACCA 7320
Db |||||
QY 2347 TCGCCAAATGTGAGCATGTCTTCTACAAATGACCTTGTCAATGGCAACCACTGCAAAACGA 2406
Db |||||
QY 7321 TCGCCAAATGTGAGCATGTCTTCTACAAATGACCTTGTCAATGGCAACCACTGCAAAACGA 7380
Db |||||
QY 2407 GGGTGGAGAACAGTGTCTCATCACTGCGACGCAAGGACCCACTCAACCGGAGAGACCA 2466
Db |||||
QY 7381 GGGTGGAGAACAGTGTCTCATCACTGCGACGCAAGGACCCACTCAACCGGAGAGACCA 7440
Db |||||
QY 2467 GGTGTGTGTGAGAGAGCCATCTGCACTGGGACATCTTTGCTTCCCAATGGGATCATTC 2526
Db |||||
QY 7441 GGTGTGTGTGAGAGAGCCATCTGCACTGGGACATCTTTGCTTCCCAATGGGATCATTC 7500
Db |||||
QY 2527 ATGTCAATTTCCAGGCCCTTTAAAGCACCCCTGCCCCGCTGACCTTTGACCCACACTGGCT 2586
Db |||||
QY 7501 ATGTCAATTTCCAGGCCCTTTAAAGCACCCCTGCCCCGCTGACCTTTGACCCACACTGGCT 7560
Db |||||
QY 2587 TGGAGACAGGATCTTCTTTGCGATCATCTGTGTGACCTGGGGCTGTGGCTTGGCTGCTT 2646
Db |||||
QY 7561 TGGAGACAGGATCTTCTTTGCGATCATCTGTGTGACCTGGGGCTGTGGCTTGGCTGCTT 7620
Db |||||
QY 2647 ACTCCTACTTTCCGATTAACCGGAGAACAAATCGGCTTCCAGCATTTTGTGAGTCGGAAGAGG 2706
Db |||||
QY 7621 ACTCCTACTTTCCGATTAACCGGAGAACAAATCGGCTTCCAGCATTTTGTGAGTCGGAAGAGG 7680
Db |||||
QY 2707 ACATTAATTTGTCAGCTCTTTGGCAAGCAGCAGCTGAGAAATATCTCGAAACCCCTTTGATG 2766
Db |||||
QY 7681 ACATTAATTTGTCAGCTCTTTGGCAAGCAGCAGCTGAGAAATATCTCGAAACCCCTTTGATG 7740
Db |||||
QY 2767 AGACCAACCTCAGCTCCCGCAGAACCTTCTTACGACCCCTTCCAGGACTCTGAAAGAC 2826
Db |||||
QY 7741 AGACCAACCTCAGCTCCCGCAGAACCTTCTTACGACCCCTTCCAGGACTCTGAAAGAC 7800
Db |||||
QY 2827 GGCAAGTTGAGGGCAATGACCCCTTGGAGACACTGTGAGGGCTGCGAGCGGAGATGCCAG 2886
Db |||||
QY 7801 GGCAAGTTGAGGGCAATGACCCCTTGGAGACACTGTGAGGGCTGCGAGCGGAGATGCCAG 7860
Db |||||
QY 2887 CCATCACTCACTGCGACCTGGGCCATCAATGTGAAATTTCTCAGCACCACTTTCCTTTAG 2946
Db |||||
QY 7861 CCATCACTCACTGCGACCTGGGCCATCAATGTGAAATTTCTCAGCACCACTTTCCTTTAG 7920
Db |||||
QY 2947 GAACGTAAGTCTCTTTAAGCATCTCAGAAAGCATACCTCATCTCTCTGGCTGATCTGGGG 3006
Db |||||
QY 7921 GAACGTAAGTCTCTTTAAGCATCTCAGAAAGCATACCTCATCTCTCTGGCTGATCTGGGG 7980
Db |||||
QY 3007 TTGTTTCTGTGGGTGAGAGATGTGTGCTGTGCCACCCAGTACAGCTTCTCTCTCTGAC 3066
Db |||||
QY 7981 TTGTTTCTGTGGGTGAGAGATGTGTGCTGTGCCACCCAGTACAGCTTCTCTCTCTGAC 8040
Db |||||
QY 3067 CTTTGGCTCTCTCTCTCTTTTGTACTCTTTCAGCTGGCACCTGCTCCATTTCTGCCCTACATG 3126
Db |||||
QY 8041 CTTTGGCTCTCTCTCTCTTTTGTACTCTTTCAGCTGGCACCTGCTCCATTTCTGCCCTACATG 8100
Db |||||
QY 3127 ATGGGTAACGTGTGATCTTTCTTCCCTGTGATGTTGAAGCTCTGNTCTTTGTATFCCAGC 3186
Db |||||

Db 8101 ATGGGTAACGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCGCTTTGTATCCACG 8160
Qy 3187 CCCTAGCCAGTGCCTGACACAGGAACTGTCACAAATAAGGTTTATGGAACAG 3240
Db 8161 CCCTAGCCAGTGCCTGACACAGGAACTGTCACAAATAAGGTTTATGGAACAG 8214

RESULT 8
LOCUS AY227444 4492 bp mRNA linear PRI 23-APR-2003
DEFINITION Homo sapiens hyaluronan receptor for endocytosis precursor (HARE)
ACCESSION AY227444
VERSION AY227444.1 GI:28848629
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4492)
AUTHORS Zhou, B., McGary, C.T., Weigel, J.A., Saxena, A. and Weigel, P.H.
TITLE Purification and molecular identification of the human hyaluronan
receptor for endocytosis
JOURNAL Glycobiology 13 (5), 339-349 (2003)
MEDLINE 22588541
PUBMED 12626425
REFERENCE 2 (bases 1 to 4492)
AUTHORS Weigel, P.H.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2003) Biochemistry and Molecular Biology,
University of Oklahoma Health Sciences Center, BMSB Room 860,
Oklahoma City, OK 73104, USA
Location/Qualifiers

FEATURES
source 1..4492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene <1..4492
/gene="HARE"
CDS <1..4251
/gene="HARE"
/codon_start=1
/product="hyaluronan receptor for endocytosis precursor"
/protein_id="AA039681.1"
/db_xref="GI:28848630"
/translation="SLPMLLEQMPDYSIFRGVLIQVNLNLALEADAYTVFAPNN
NAENIIEKKVLSLEEDVLRVHVLEKLLANDLHNGHRETM LGFYFLFLHND
QLYVNEAPINTVNTADKGVHGLKVLIEIKRNDNDTIIIRGRCTCSELTPFP
GTSKLNKRCIYTSYFMGRRTLFIGQPKVTVITRECCAGFPQCPQCNQAQ
NVCFNGICLDGWNVTGVCCEGSGTACETCTRGKYIHCDQACSCVHGRNGPL
GDGCDVWGRGVHCDNATTEDCNGTCHTSANCLTNSDGTASCKCAAGFGNGTIC
TAINACEISNGGCSAKADKRTTPGRVCTCKAGYTGDIIVCLEINPCLENHGGCDKN
AECTQTGNQAACNCLPAYTGKGVCTLIYVCLTKNGGCSFPAICNHGQVETCTCK
ENYIGDFTCRGSIYQELPNKTSQYFQJQHEFVKDVLVGPPTVPAPUSAFDEE
ARKWDKYGMLPQVLRVHVACHQLLENLKIISNATSLQGEPIVISVSQSTVIN
KAKITSSDIISNIGVHVIDKLLSPKLLITPKNSGRILQNLTLATNGYIKFSNL
IDKSLVLPSTDPHTPVTLPWPTQALHALPAEQDPLFNODNKKLKEYLKPVR
TADVLADLPSTKPTLQGSSELSVKCGAGRDIGLFLNGOTCRIVQBELLPDLGVAY
GIDCLLIDPTLGGRCDDTTTFDASEGSCSVNTPSCRPSKPKVKQCLNLPKRN
LEGRCERSLVQIIPRCCKGYFRDQCPCPGPDAPCNRRGVCLDQIYATGECKNTG
FNGTACEMWPFGRFDCPLPCGSDHQQDDGITGSGQCLCEGTGTGSPCDTQAVLPA
VCTPPCSAHATCKENNTCECLNDYEGDGTCTVVPDFCKDNGGCKARVACSQKGTKVS
CSCQKGYGDHSCETIDPCADLNGGCHHEATCKMTGPGHKCECKSHYVGDGNCE
PQPLDPLRQDNGCHADKCVLDHFDQTVVGFHLRSLPGYKLTIPKAREACANE
ALMTATYNLQSAQKAKVLCASGMLGTGRAVPTAPASONGSCGVGVVDYGPENK
SEMDVFCYRMKDVNCTCKVYVGVGFSCSNLLQVLMSPSLTNFLTEVLAYSNSA
RGRLEHLTDLUSIRGLTFVPQNSGLNETUSGRDIEHHLANSMFFYNDLVNGTTL
QTRLSKLITASHQDPLQPTETFRVDRALQMDIFASNGIITHVSRPKAPPAPVTL
THTGLCAGTFFAIILVTGAVALAAYSFRINRTICFQHFSEEDINVAALQKQOPEN
ISNPLYESTTSAPPSYDPFTDSEERQLEGNDRLLT"
mat_peptide 1..4248
/gene="HARE"
/product="190 kDa hyaluronan receptor for endocytosis"

/function="mediates endocytosis of hyaluronan or
chondroitin sulfate via the clathrin coated pit pathway"
/note="generated by proteolytic processing of a larger
protein; probably a product of the HARE/Stablin-2 gene;
190 kDa protein purified from spleen"
ORIGIN
Query Match 94.9%; Score 3094.2; DB 9; Length 4492;
Best Local Similarity 99.9%; Pred No. 0;
Matches 3096; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 7 AGCAGGAGCTTCCCAGAACCCGAAACTTCCAGTATTTCTTCAGTTGCAAGGAGCAAT 66
Db 1394 ATCAGGAGCTTCCCAGAACCCGAAACTTCCAGTATTTCTTCAGTTGCAAGGAGCAAT 1453
Qy 67 TCGTGAAGAATCTGGTCGGCCAGCCCTTCTACCTGTTTTCACCTTATCTGCGAGCT 126
Db 1454 TCGTGAAGAATCTGGTCGGCCAGCCCTTCTACCTGTTTTCACCTTATCTGCGAGCT 1513
Qy 127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCAAGGTTCTTC 186
Db 1514 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCAAGGTTCTTC 1573
Qy 187 GGTACCATGTGGTGCCTTGGCCACCAAGCTGCTTCTGGAAAACTGAAATTTGATCTCAAAATG 246
Db 1574 GGTACCATGTGGTGCCTTGGCCACCAAGCTGCTTCTGGAAAACTGAAATTTGATCTCAAAATG 1633
Qy 247 CTACTTCCCTCCAAGGAGGACCAATAGTCACTCCGCTCTCTCAGAGCAGGTTGATATAA 306
Db 1634 CTACTTCCCTCCAAGGAGGACCAATAGTCACTCCGCTCTCTCAGAGCAGGTTGATATAA 1693
Qy 307 ATAATAAGGCTAAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTTGTTCATATCA 366
Db 1694 ACAATAAGGCTAAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTTGTTCATATCA 1753
Qy 367 TAGACAAATGCTATCTCCCAAAAATTTGCTTTATCACTCCCAAGACAACTCTGGAAGAA 426
Db 1754 TAGACAAATGCTATCTCCCAAAAATTTGCTTTATCACTCCCAAGACAACTCTGGAAGAA 1813
Qy 427 TTCTGCAAAATCTTACGACTTTGGCAACAACAATGGCTACATCAAAATTTAGCACTTAA 486
Db 1814 TTCTGCAAAATCTTACGACTTTGGCAACAACAATGGCTACATCAAAATTTAGCACTTAA 1873
Qy 487 TACAGGACTCAGGTTTGTGTGAGTGTATCACCAGTCCCATCCACACCCCAAGTCACTCTCT 546
Db 1874 TACAGGACTCAGGTTTGTGTGAGTGTATCACCAGTCCCATCCACACCCCAAGTCACTCTCT 1933
Qy 547 TCTGGCCACCGACCAAGCCCTCCATGCCCTACCTGCTGGAACAACAGGACTTCTGTTTCA 606
Db 1934 TCTGGCCACCGACCAAGCCCTCCATGCCCTACCTGCTGGAACAACAGGACTTCTGTTTCA 1993
Qy 607 ACCAAGACAAACAGGACCAAGCTGAAGGAGTATTTTGAAGTTTTCATGTGATACGAGATGCCA 666
Db 1994 ACCAAGACAAACAGGACCAAGCTGAAGGAGTATTTTGAAGTTTTCATGTGATACGAGATGCCA 2053
Qy 667 AGGTTTATAGTCTGGATCTTCCACATCCATCCCTGCTGGAAGACCCCTGCAAGGTTTCAGAGC 726
Db 2054 AGGTTTATAGTCTGGATCTTCCACATCCATCCCTGCTGGAAGACCCCTGCAAGGTTTCAGAGC 2113
Qy 727 TGAGTGTGAATGTGGAGCTGCGAGGACATCGGTGACCTCTTTCTGATGCGCAAACT 786
Db 2114 TGAGTGTGAATGTGGAGCTGCGAGGACATCGGTGACCTCTTTCTGATGCGCAAACT 2173
Qy 787 GCAGAAATGTGCGAGCGGAGCTCTTGTGTTGACCTGGGTGTGGCCCTACGCGATTGACTGTC 846
Db 2174 GCAGAAATGTGCGAGCGGAGCTCTTGTGTTGACCTGGGTGTGGCCCTACGCGATTGACTGTC 2233
Qy 847 TCGTGATTGATCCCACTTGGGGCGCGCTGTGACACCTTTACTACTTTTCGATGCCTCGG 906
Db 2234 TCGTGATTGATCCCACTTGGGGCGCGCTGTGACACCTTTACTACTTTTCGATGCCTCGG 2293
Qy 907 GGGAGTGTGGGAGCTGTGCTCAATATCTCCAGCTGCCAAGGTGGAGTAACCAAGGGTGG 966

Db 2294 GGGAGTGGGAGCTGTGTAATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGAGGTG 2353
Qy 967 TGAAGCAAGTGTCTTACAACTGCCCCCTTCAAGAGGAACTGGAAGGCTGCCGGGAGC 1026
Db 2354 TGAAGCAAGTGTCTTACAACTGCCCCCTTCAAGAGGAACTGGAAGGCTGCCGGGAGC 2413
Qy 1027 GGTGAGCCTGGTATACAGATCCCAAGTGTCTCAAGGCTTCTCGGGGAGACTGTC 1086
Db 2414 GGTGAGCCTGGTATACAGATCCCAAGTGTCTCAAGGCTTCTCGGGGAGACTGTC 2473
Qy 1087 AGGCTGCGCTGGAGGACAGATGCCGCTGTAAATAACCGGGTGTCTGCTTGTATCAGT 1146
Db 2474 AGGCTGCGCTGGAGGACAGATGCCGCTGTAAATAACCGGGTGTCTGCTTGTATCAGT 2533
Qy 1147 ACTGCGCACCGGAGGTGTAATGCAACCGGCTTCAATGGGACGGCTGTGAGATGT 1206
Db 2534 ACTGCGCACCGGAGGTGTAATGCAACCGGCTTCAATGGGACGGCTGTGAGATGT 2593
Qy 1207 GCTGGCCGGGAGATTCCGGCTGTGATGTCTGCCCCGTGTCTGCTGCTGCTGCTGCTGCT 1266
Db 2594 GCTGGCCGGGAGATTCCGGCTGTGATGTCTGCCCCGTGTCTGCTGCTGCTGCTGCTGCT 2653
Qy 1267 GCGATGATGGCATCACCGGCTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1326
Db 2654 GCGATGATGGCATCACCGGCTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2713
Qy 1327 CGTGACATCAGGCACTTTTGGCTGCGAGTGTGTACCGCTTCTTGTCTGCTCATGCCA 1386
Db 2714 CGTGACATCAGGCACTTTTGGCTGCGAGTGTGTACCGCTTCTTGTCTGCTCATGCCA 2773
Qy 1387 CCTGTAGGAAACAAACAGTGTGATGTAACTCGGATTTAGAGGTGAACGGAATCAAT 1446
Db 2774 CCTGTAGGAAACAAACAGTGTGATGTAACTCGGATTTAGAGGTGAACGGAATCAAT 2833
Qy 1447 GCACAGTGTGATTTCTGCAAAACAGGACACGGGGCTGTGCAAAAGGTGCCAGATGCT 1506
Db 2834 GCACAGTGTGATTTCTGCAAAACAGGACACGGGGCTGTGCAAAAGGTGCCAGATGCT 2893
Qy 1507 CCCAGAGGGCACGAAGTCTCTGCAAGTGTGCAAGAGGGATACAAAGGGGACGGGCACA 1566
Db 2894 CCCAGAGGGCACGAAGTCTCTGCAAGTGTGCAAGAGGGATACAAAGGGGACGGGCACA 2953
Qy 1567 GCTGCACAGATAGACCCCTGTGCAGACGGCTTACGGAGGTGTACAGACACGCCA 1626
Db 2954 GCTGCACAGATAGACCCCTGTGCAGACGGCTTACGGAGGTGTACAGACACGCCA 3013
Qy 1627 CCTGTAGATGACAGGCCCCGGGCAAGTGTGAGTGTAAAGTCACTATGTGCGAG 1686
Db 3014 CCTGTAGATGACAGGCCCCGGGCAAGTGTGAGTGTAAAGTCACTATGTGCGAG 3073
Qy 1687 ATGGGCTGAATGTGAGCCGAGCAGTGTGCCATTTGACCCGTGTCTTACAGGACAAATGGGC 1746
Db 3074 ATGGGCTGAATGTGAGCCGAGCAGTGTGCCATTTGACCCGTGTCTTACAGGACAAATGGGC 3133
Qy 1747 AGTCCCATGACAGCCCAATGTGTGACCTTCCACTTCCAGATACCACTGTGGGTGT 1806
Db 3134 AGTCCCATGACAGCCCAATGTGTGACCTTCCACTTCCAGATACCACTGTGGGTGT 3193
Qy 1807 TCCATCTACGCTCCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCT 1866
Db 3194 TCCATCTACGCTCCCCACTGGGCCAGTATAGCTGACCTTTGACAAAGCCAGAGAGGCT 3253
Qy 1867 GTGCCAAGAGCTGCCACCTAGGCAACCTTACAAACAGCTTCTCTATGCCAGAGGCCA 1926
Db 3254 GTGCCAAGAGCTGCCACCTAGGCAACCTTACAAACAGCTTCTCTATGCCAGAGGCCA 3313
Qy 1927 AGTACCACCTGTCTCAGCAGCTGTGAGAGACGGGGCGGTGTGCTTACCCCAAGCCT 1986
Db 3314 AGTACCACCTGTCTCAGCAGCTGTGAGAGACGGGGCGGTGTGCTTACCCCAAGCCT 3373
Qy 1987 TCGCTCCCAAGACTGTGGCTCTGCTGTGCTGGATGAGTATGAGCTATGAGCTAGACCCA 2046
Db 3374 TCGCTCCCAAGACTGTGGCTCTGCTGTGCTGGATGAGTATGAGCTATGAGCTAGACCCA 3433

Qy 2047 ACAAGAGTGAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACCTGCACTGCA 2106
Db 3434 ACAAGAGTGAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAACCTGCACTGCA 3493
Qy 2107 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAAGTGGGAACTGTGTCAGGTCTGTATGT 2166
Db 3494 AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAAGTGGGAACTGTGTCAGGTCTGTATGT 3553
Qy 2167 CTTTCCCTCACTCACAACTTCTGACGGAAGTGTGCTGCTATTTCCAAACAGCTCAGTTC 2226
Db 3554 CTTTCCCTCACTCACAACTTCTGACGGAAGTGTGCTGCTATTTCCAAACAGCTCAGTTC 3613
Qy 2227 GAGGCGGTGCAATTTCTAGAACACCTGACTGACCTGTCTCATCGGGGACCCCTTTTGTGC 2286
Db 3614 GAGGCGGTGCAATTTCTAGAACACCTGACTGACCTGTCTCATCGGGGACCCCTTTTGTGC 3673
Qy 2287 CACAGACAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGGGAGCAATCGAGACACACC 2346
Db 3674 CACAGACAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGGGAGCAATCGAGACACACC 3733
Qy 2347 TCGCAATGTGAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACTCTGCAAAACGA 2406
Db 3734 TCGCAATGTGAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACTCTGCAAAACGA 3793
Qy 2407 GGGTGGGAAGCAAGTGTCTCATCTGCGACAGGAGCCCACTCCAAACCGACGGAGACCA 2466
Db 3794 GGGTGGGAAGCAAGTGTCTCATCTGCGACAGGAGCCCACTCCAAACCGACGGAGACCA 3853
Qy 2467 GGTGTGTGATGGAAGACCAATCTGCAAGTGGGACATCTTTTGTCTCAATGGGATCATTC 2526
Db 3854 GGTGTGTGATGGAAGACCAATCTGCAAGTGGGACATCTTTTGTCTCAATGGGATCATTC 3913
Qy 2527 ATGTCAATTTCCAGGCTTTTAAAGCACCCCTGCCCCCGTGACCTTGACCCACACTGGCT 2586
Db 3914 ATGTCAATTTCCAGGCTTTTAAAGCACCCCTGCCCCCGTGACCTTGACCCACACTGGCT 3973
Qy 2587 TGGGAGCAGGATCTTCTTTGCCATCATCTCTGTGTGACTGGGGCTGTGTGCTTGGCTGT 2646
Db 3974 TGGGAGCAGGATCTTCTTTGCCATCATCTCTGTGTGACTGGGGCTGTGTGCTTGGCTGT 4033
Qy 2647 ACTCCTACTTTGGATTAACCGGAGAACATTCGGCTTTCAGCATTTTGTAGTCGGAAGAGG 2706
Db 4034 ACTCCTACTTTGGATTAACCGGAGAACATTCGGCTTTCAGCATTTTGTAGTCGGAAGAGG 4093
Qy 2707 ACATTAATGTGAGCTTCTTGGCAAGCAGCCTGAGAAATATCTCGAAACCCCTTGTATG 2766
Db 4094 ACATTAATGTGAGCTTCTTGGCAAGCAGCCTGAGAAATATCTCGAAACCCCTTGTATG 4153
Qy 2767 AGACCAACCTCAGCTCCCCCAGAACCTTCTTACGACCCCTTACGAGCTTCTGAAGAAC 2826
Db 4154 AGACCAACCTCAGCTCCCCCAGAACCTTCTTACGACCCCTTACGAGCTTCTGAAGAAC 4213
Qy 2827 GGCAGCTTGGGGCAATGACCCCTTGAAGACATCTGTCAGGGCTTGGACGGGAGATGCCAG 2886
Db 4214 GGCAGCTTGGGGCAATGACCCCTTGAAGACATCTGTCAGGGCTTGGACGGGAGATGCCAG 4273
Qy 2887 CCATCACTCACTGCCCATCAACTGTGAATTTCTCAGCACAGTTCCTTTTAG 2946
Db 4274 CCATCACTCACTGCCCATCAACTGTGAATTTCTCAGCACAGTTCCTTTTAG 4333
Qy 2947 GAACGTAAAGTCTTTTAAAGCACTCAGAAAGCATACCTCATCTCTGGCTGATCTGGGG 3006
Db 4334 GAACGTAAAGTCTTTTAAAGCACTCAGAAAGCATACCTCATCTCTGGCTGATCTGGGG 4393
Qy 3007 TTGTTTCTGTGGTGAAGATGTGCTGTGCCCAACCCAGTACAGTTCCTCTCTGAC 3066
Db 4394 TTGTTTCTGTGGTGAAGATGTGCTGTGCCCAACCCAGTACAGTTCCTCTCTGAC 4453
Qy 3067 CCTTTGGCTCTTCTTCTTGTACTCTTTCAGCTGGCAC 3105
Db 4454 CCTTTGGCTCTTCTTCTTGTACTCTTTCAGCTGGCAC 4492

RESULT 9
LOCUS HSM801377 3642 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFpZp434E0321 (from clone DKFpZp434E0321).
ACCESSION AL133021
VERSION AL133021.1 GI:6453577
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3642)
AUTHORS Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-1999) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY
COMMENT This clone (DKFpZp434E0321) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES

location/Qualifiers
1..3642
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFpZp434E0321"
/tissue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
join(7..2607,2609..3217)
/gene="DKFpZp434E0321"
join(<7..2607,2609..3217)
/note="DKFpZp434E0321"
/gene="strong similarity to KIAA0246"
The frame shift was determined manually"
/codon_start=1
/product="hypothetical protein"
/protein_id="CB61358.1"
/db_xref="GI:6453578"
/db_xref="GOA:Q9UF98"
/db_xref="SPTREMBL:Q9UF98"
/translation="VGEAVGTASCKCAAGFQNGTICTAINACEISNGGCSAKADCKR
TTPGRVCTCKAGYTGDIIVCLEINPLENHGGCDKNAECTOTGPNQACNCLPAYTG
DGKVTLLINVLTKGGCSEFAICNHTGOVETCTCKNYIGDGTGRGSIYQELPKN
PKISQIFQQLHFVNDLVGPGFTVFPAPLSAFAEBEARVKDWKYGMLPQVLRHV
ACHQLLENKLISNATSLQGPFIIVISQSVIYNNKAKIISDIIISTNGIVHIDK
WPTDQALHALPAEQDFPNQDNKDKLEYLFKFLVIRDAKVLAVDLPSTANKTLQGS
ELSVKCGACGRDIFLNGQTCRIIVQRELLFDLVGAYGTDCLLDIPTLGGRCDTFTF
DASGCGSVNTPSPCRWSKPKGVKQCLNLPKFNLEGGERLSLVQIPRCCKGY
FQDCQACPGGDFPCNRRNGVCLDQISATGECKNTGNFNTACENCMWGRFGPDLPC
GCSHGQDDGTIVGSGQCLICETGTPGSCDQAVLPVCTPPCSAHATCKENITCECN
LDEGSDGTIVVDFCKQDNGCAKRVSCQKGYKGDGSHSCNTEIDPCA
DLNGCKEHAETCKWTGPKGKCEKSHVYVGLNCEPEQLPDLCLQDNGCHADAK
CVDLHFQDFTVGHLSPLGVKLTDFDKAREACANEATMATYNQLSVAOKAKVHLC
SAGWLETGVAVPTAFASQNGSGVVGIVDYGPRNPKSEMDVFCYRMKGSAGLPQOL
SRPCLSRTPDLSIRGLTFVQNSGLGENETLSGRDIEHHLANVSMFPYNDLVNGTT
LQTRGLSGLIITFAILLVGAVALAAYSTPRINRTIGFQHFESSEEDINVAALGKQOPE
LNTSFLYESTTSAPPPEPSYDPTDSEERQLEGNDPLRTI"
3574..3579
3603

polyA_signal
polyA_site
ORIGIN

Query Match

92.6%; Score 3017.8; DB 9; Length 3642;

Best Local Similarity 96.5%; Pred. No. 0;
Matches 3139; Conservative 0; Mismatches 8; Indels 107; Gaps 1;
QY 7 AGCAGGAGCTTCCCAAGAACCCGAAACATTTCCAGTATTCTTCCAGTTGCAGGAGCAT 66
DB 467 ATCAGGAGCTTCCCAAGAACCCGAAACATTTCCAGTATTCTTCCAGTTGCAGGAGCAT 526
QY 67 TCGTGAAGAAGCTCGGTGGGCGCCAGGCGCTTCACTGTGTTTTGACATTTATCTGAGCGCT 126
DB 527 TCGTGAAGAAGCTCGGTGGGCGCCAGGCGCTTCACTGTGTTTTGACATTTATCTGAGCGCT 586
QY 127 TTGATGAGGAAGCTCGGTGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 186
DB 587 TTGATGAGGAAGCTCGGTGTTAAAGACTGGGACAAATACGGTTTAATGCCCCAGGTTCTTC 646
QY 187 GGTACCATGTGGTGGCGCTGCCACAGCTGCTTCTGGAAAACTGAAATGATGATCAAAATG 246
DB 647 GGTACCATGTGGTGGCGCTGCCACAGCTGCTTCTGGAAAACTGAAATGATGATCAAAATG 706
QY 247 CTACTTCCCTCAAGAGAGAGCAATAGTCACTCTCCGTCTCTCAGAGCAGCGGTGTATATAA 306
DB 707 CTACTTCCCTCAAGAGAGAGCAATAGTCACTCTCCGTCTCTCAGAGCAGCGGTGTATATAA 766
QY 307 ATATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGATTTGTTCATATCA 366
DB 767 ATATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGATTTGTTCATATCA 826
QY 367 TAGCAAAATGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGAA 426
DB 827 TAGCAAAATGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGAA 886
QY 427 TTCTCAAAATCTTACGACTTTTGGCAACAAAACAATGGGTACATCAAAATTTAGCAACTTAA 486
DB 887 TTCTGCAAAATCTTACGACTTTTGGCAACAAAACAATGGGTACATCAAAATTTAGCAACTTAA 946
QY 487 TACAGGACTCAGGTTTGTGTAGTGTCATCACCGATCCCATCCACACCCAGTCACTCTCT 546
DB 947 TACAGGACTCAGGTTTGTGTAGTGTCATCACCGATCCCATCCACACCCAGTCACTCTCT 1006
QY 547 TCTGGCCACACGACCAAGCCCTCCATGCCCTACCTGTGAAACAAGAGGACTTCTGTGTCA 606
DB 1007 TCTGGCCACACGACCAAGCCCTCCATGCCCTACCTGTGAAACAAGAGGACTTCTGTGTCA 1066
QY 607 ACCAAGACAAACAAGACAAAGCTGAAGAGGATTTTGAAGTTTTCATGTATACGAGATGCCA 666
DB 1067 ACCAAGACAAACAAGACAAAGCTGAAGAGGATTTTGAAGTTTTCATGTATACGAGATGCCA 1126
QY 667 AGTTTGTAGCTGTGATCTTCCACATCCATCGCTGGAAGACCCCTGCAAGGTTTCAGAGC 726
DB 1127 AGTTTGTAGCTGTGATCTTCCACATCCATCGCTGGAAGACCCCTGCAAGGTTTCAGAGC 1186
QY 727 TGAGTGTGAATGTGAGCTGGCAGGACATCGGTGACCTCTTCTGAAATGGCCAAACCT 786
DB 1187 TGAGTGTGAATGTGAGCTGGCAGGACATCGGTGACCTCTTCTGAAATGGCCAAACCT 1246
QY 787 GCAGAAATGTGACGGGAGCTCTTGTGTTGACCTGGGTGGCCTACGCAATGACTGTC 846
DB 1247 GCAGAAATGTGACGGGAGCTCTTGTGTTGACCTGGGTGGCCTACGCAATGACTGTC 1306
QY 847 TGCTGATTGATCCCACTCTGGGGGGCGCTGTGACACCTTTACTATTTCGATGCTCGG 906
DB 1307 TGCTGATTGATCCCACTCTGGGGGGCGCTGTGACACCTTTACTATTTCGATGCTCGG 1366
QY 907 GGGAGTGTGGGAGCTGTGTAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGGTG 966
DB 1367 GGGAGTGTGGGAGCTGTGTAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGGTG 1426
QY 967 TGAAGCAGAGGAGTGTCTTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCGGGAGC 1026
DB 1427 TGAAGCAGAGGAGTGTCTTACAACCTGCCCTTCAAGAGGAACCTGGAAGGCTGCGGGAGC 1486
QY 1027 GGTGAGCTGTGTGATACAGATCCCAGGCTGCAAGGCTACTTCCGGCGGAGCTGTC 1086

Db 1487 GGTGCAGCCTGGTGATACAGATCCCGCAGGTGCTGCAAGGGCTACTTTCGGGCGAGACTGTC 1546
QY 1087 AGGCTTGCCCTGGAGGACAGATGCCCGGTGTAATAACCGGGGTGTCTGCTTGTATCAGT 1146
Db 1547 AGGCTTGCCCTGGAGGACAGATGCCCGGTGTAATAACCGGGGTGTCTGCTTGTATCAGT 1606
QY 1147 ACTCGGCCACCGGAGAGTGAATGCAACACCGGCTTCAATGGGAGCGGTGTGATGAT 1206
Db 1607 ACTCGGCCACCGGAGAGTGAATGCAACACCGGCTTCAATGGGAGCGGTGTGATGAT 1666
QY 1207 GCTGGCCGGGAGATTTCGGGCTGATTTGTCTGCCCTGTGGCTGCTCAGACACGACAGT 1266
Db 1667 GCTGGCCGGGAGATTTCGGGCTGATTTGTCTGCCCTGTGGCTGCTCAGACACGACAGT 1726
QY 1267 GCGATGATGGGATCAAGGCTCCGGGAGTGCCTCTGTGAAACCGGGGTGGACAGGCCCT 1326
Db 1727 GCGATGATGGGATCAAGGCTCCGGGAGTGCCTCTGTGAAACCGGGGTGGACAGGCCCT 1786
QY 1327 CGTGTGACACTCAGGACGTTTTGCTCGAGTGTGTACGCCCTCTGTTCTGCTCATGCCA 1386
Db 1787 CGTGTGACACTCAGGACGTTTTGCTCGAGTGTGTACGCCCTCTGTTCTGCTCATGCCA 1846
QY 1387 CCTGTAAAGGAACAACAGTGTAGTGTAACTGGAATTATGAAGTGTACGGAATCACAT 1446
Db 1847 CCTGTAAAGGAACAACAGTGTAGTGTAACTGGAATTATGAAGTGTACGGAATCACAT 1906
QY 1447 GCACAGTTGTGATTTCTGCAAAACAGGACACCGGGGTGTGCAAGGTGCCAGATGCT 1506
Db 1907 GCACAGTTGTGATTTCTGCAAAACAGGACACCGGGGTGTGCAAGGTGCCAGATGCT 1966
QY 1507 CCCAGAGGGCACGAAGTCTCTGACGTGCCAGAGGGATACAAAGGACCGGCACA 1566
Db 1967 CCCAGAGGGCACGAAGTCTCTGACGTGCCAGAGGGATACAAAGGACCGGCACA 2026
QY 1567 GCTGCACAGATAGACCCCTGTGCAGACGGCTTTAAACGGAGGTGTCAACAGACAGCCA 1626
Db 2027 GCTGCACAGATAGACCCCTGTGCAGACGGCTTTAAACGGAGGTGTCAACAGACAGCCA 2086
QY 1627 CCTGTAGATGACAGGCCCGGGCAAGCAACAGTGTAGTGTAAAGTCACTATGTCCGAG 1686
Db 2087 CCTGTAGATGACAGGCCCGGGCAAGCAACAGTGTAGTGTAAAGTCACTATGTCCGAG 2146
QY 1687 ATGGGCTGAACGTGTGAGCCGAGCAGCTGCCATGACCGCTGTACAGACAAATGGGC 1746
Db 2147 ATGGGCTGAACGTGTGAGCCGAGCAGCTGCCATGACCGCTGTACAGACAAATGGGC 2206
QY 1747 AGTGCATGACAGCCGCAAAATGTGTGCAACCTCCACTCCAGGATACCACTGTGGGGTGT 1806
Db 2207 AGTGCATGACAGCCGCAAAATGTGTGCAACCTCCACTCCAGGATACCACTGTGGGGTGT 2266
QY 1807 TCCATCTACGCTCCCACTGGGCCAGTATTAAGCTGACCTTTGACAAAGCCAGAGGCTT 1866
Db 2267 TCCATCTACGCTCCCACTGGGCCAGTATTAAGCTGACCTTTGACAAAGCCAGAGGCTT 2326
QY 1867 GTGCCAACAGAGCTGGACCACTGCAACCTACAACAGCTCTCTATGSCCCAGAGGCCA 1926
Db 2327 GTGCCAACAGAGCTGGCCNACTGCAACCTACAACAGCTCTCTATGSCCCAGAGGCCA 2386
QY 1927 AGTACCACCTGTCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTCCTTACCCACAGCCT 1986
Db 2387 AGTACCACCTGTCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTCCTTACCCACAGCCT 2446
QY 1987 TCGGCTCCAGAACTGTGGCTCTGGTGTGGTGGGATAGTGGATATAGACCTTAGACCCA 2046
Db 2447 TCGGCTCCAGAACTGTGGCTCTGGTGTGGTGGGATAGTGGATATAGACCTTAGACCCA 2506
QY 2047 ACAAGAGTGAATGTGGAGTGTCTTCTGCTATCGGATGAAGATGTGAATGCACTGCA 2106
Db 2507 ACAAGAGTGAATGTGGAGTGTCTTCTGCTATCGGATGAAGATGTGAATGCACTGCA 2547
QY 2107 AGTGGGCTATGTGGGAGATGGCTTCTCATGCACTGGGGAACCTGCTGCGAGTCTGATGT 2166
Db 2548 ----- 2547

QY 2167 CCTTCCCTCACTCAAAACTTCTCTGACGAGAGTGTGCGCTATTTCACACAGCTCAGTTC 2226
Db 2548 -----GGAAGTGTGGCTATTTCACACAGCTCAGTTC 2579
QY 2227 GAGGCGGTGCAATTTCTAGAAACACTGACTGACCTGTCCATCCGCGGACCCCTTTGTGC 2286
Db 2580 GAGGCGGTGCAATTTCTAGAAACACTGACTGACCTGTCCATCCGCGGACCCCTTTGTGC 2639
QY 2287 CACAGAACAGTGGGCTGGGAGATGAGACCTTGTCTGGGCGGACATCGAGCACACC 2346
Db 2640 CACAGAACAGTGGGCTGGGAGATGAGACCTTGTCTGGGCGGACATCGAGCACACC 2699
QY 2347 TCGCAATGTGAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACTTGCACAAACA 2406
Db 2700 TCGCAATGTGAGCATGTTTTTCTACAATGACCTTGTCAATGGCACCACTTGCACAAACA 2759
QY 2407 GGGTGGGAAGCAAGTGTCTCATCTGACGACGAGACCCACTCTCAAACGAGCGGAGACCA 2466
Db 2760 GGGTGGGAAGCAAGTGTCTCATCTGACGACGAGACCCACTCTCAAACGAGCGGAGACCA 2819
QY 2467 GGTGTTGTTGATGGAAGAGCCATTTCTGCACTGGGACATCTTTGCCCTCAATGGGATCATT 2526
Db 2820 GGTGTTGTTGATGGAAGAGCCATTTCTGCACTGGGACATCTTTGCCCTCAATGGGATCATT 2879
QY 2527 ATGTCAATTTCCAGGCTTTTAAAGCACCCCTGCCCCGCTGACCTTGACCCACACTGGCT 2586
Db 2880 ATGTCAATTTCCAGGCTTTTAAAGCACCCCTGCCCCGCTGACCTTGACCCACACTGGCT 2939
QY 2587 TGGGAGCAGGATCTTCTTGGCATCATCTCTGGTGTGACTGGGGCTGTTCCTTGGCTGCTT 2646
Db 2940 TGGGAGCAGGATCTTCTTGGCATCATCTCTGGTGTGACTGGGGCTGTTCCTTGGCTGCTT 2999
QY 2647 ACTCTCACTTTCGGATAAACCGGAGAACATCGGCTTCCAGCATTTTCAGTTCGGAAGAGG 2706
Db 3000 ACTCTCACTTTCGGATAAACCGGAGAACATCGGCTTCCAGCATTTTCAGTTCGGAAGAGG 3059
QY 2707 ACATTAATGTTGACGCTCTTGGCAAGCAGCAGCTGAGAAATATCTCGAACCCCTTGTATG 2766
Db 3060 ACATTAATGTTGACGCTCTTGGCAAGCAGCAGCTGAGAAATATCTCGAACCCCTTGTATG 3119
QY 2767 AGACGACAACTCAGTCCGCCAGAACCTTCTAGACCCCTTCAOGGACTCTGAAGAAC 2826
Db 3120 AGACGACAACTCAGTCCGCCAGAACCTTCTAGACCCCTTCAOGGACTCTGAAGAAC 3179
QY 2827 GGCAGCTTGGGGCAATGACCCCTTGGAGCAGCTGAGGGCCTGAGCGGAGATGCCAG 2886
Db 3180 GGCAGCTTGGGGCAATGACCCCTTGGAGCAGCTGAGGGCCTGAGCGGAGATGCCAG 3239
QY 2887 CCATCACTCACTGCCACTGGGCCATCAACTGTGAATTTCTCAGCACCACTTGCCTTTTAG 2946
Db 3240 CCATCACTCACTGCCACTGGGCCATCAACTGTGAATTTCTCAGCACCACTTGCCTTTTAG 3299
QY 2947 GACGTAAGTCTTTTAAGCACTCAGAACCCATACCTCATCTCTCGGCTGATCTGGGG 3006
Db 3300 GACGTAAGTCTTTTAAGCACTCAGAACCCATACCTCATCTCTCGGCTGATCTGGGG 3359
QY 3007 TTGTTTCTGTGGGTGAGAGATGTTGCTGTGCCCAACCCAGTACAGCTTCTCCTCTCTGAC 3066
Db 3360 TTGTTTCTGTGGGTGAGAGATGTTGCTGTGCCCAACCCAGTACAGCTTCTCCTCTCTGAC 3419
QY 3067 CCTTGGCTCTTCTTCTTGTACTCTTCACTGGCACCTGCTCCCACTTCTGCCCCATCATG 3126
Db 3420 CCTTGGCTCTTCTTCTTGTACTCTTCACTGGCACCTGCTCCCACTTCTGCCCCATCATG 3479
QY 3127 ATGGGTAACTGTGATCTTCTTCTTGTAGTTGAAGCCTCTCTTCTTGTATCCAGC 3186
Db 3480 ATGGGTAACTGTGATCTTCTTCTTGTAGTTGAAGCCTCTCTTCTTGTATCCAGC 3539
QY 3187 CCTTAGCCAGTCTGACACAGCACTGTGCACTAAAGGTTTATGGACAGGAAAAA 3246
Db 3540 CCTTAGCCAGTCTGACACAGCACTGTGCACTAAAGGTTTATGGACAGGAAAAA 3599


```
QY 3247 AAAAAAAAAAAAA 3260
Db 3600 AGTCAAAAAAAAA 3613

RESULT 10
AX686192 AX686192 8444 bp DNA linear PAT 29-MAR-2003
LOCUS Sequence 1 from Patent WO02059315.
DEFINITION AX686192
ACCESSION AX686192
VERSION AX686192.1 GI:29372026
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 Shimkets, R.A., Patturajan, M., Vernet, C.A., Casman, S.J.,
Malyankar, U., Shenoy, S., Spytek, K.A., Gangolli, E., Miller, C.,
Boldog, F., Li, L., Taupier, R.J., Kekuda, R., Smithson, G.,
Zerhusen, B.D., Liu, X., Colman, S.D., Tchernev, V., Si, J., Edinger, S.,
Stone, D., Sciore, P., Millet, I., and Rothenberg, M.
Human nucleic acids and polypeptides and methods of use thereof
Patent: WO 02059315-A 1 01-AUG-2002;
JOURNAL Curagen Corporation (US)
FEATURES
source
1..8444
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 92.0%; Score 2999.8; DB 6; Length 8444;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 3128; Conservative 0; Mismatches 18; Indels 108; Gaps 1;

QY 7 AGCAGAGCTTCCCAAGAACCCGAAACTTCCAGTATTCTTCCAGTTTCCAGTTGCGAGGACATT 66
Db ATCAGGAGCTTCCCAAGAACCCGAAACTTCCAGTATTCTTCCAGTTTCTTCCAGTTGCGAGGACATT 5338

QY 67 TCGTGAAGATCTGTCGCGCCAGGCCCTTCACTGTTTTTGGACCTTTATCTGAGCCT 126
Db TCGTGAAGATCTGTCGCGCCAGGCCCTTCACTGTTTTTGGACCTTTATCTGAGCCT 5398

QY 127 TTGATGAGAACTCGGGTAAAGCTGGGACAAATACGGTTAAATGCGCCAGGTTCTTC 186
Db TTGATGAGAACTCGGGTAAAGCTGGGACAAATACGGTTAAATGCGCCAGGTTCTTC 5458

QY 187 GGTACCATCTGTCGCTGCCACCACTGCTTCTGGAACACCTGAAATTTGATCTCAAATG 246
Db GGTACCATCTGTCGCTGCCACCACTGCTTCTGGAACACCTGAAATTTGATCTCAAATG 5518

QY 247 CTACTTCCCTCCAGAGAGCCAAATAGTCAATCTCCGTTCTCTCAGACACGGTGTATATAA 306
Db CTACTTCCCTCCAGAGAGCCAAATAGTCAATCTCCGTTCTCTCAGACACGGTGTATATAA 5578

QY 307 ATAATAAGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGATTGTTTATATCA 366
Db ATAATAAGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGATTGTTTATATCA 5578

QY 367 TAGACAAATTTGCTATCTCCCAAAATTTGCTTATCTACTCCCAAGACAACTCTGGAAGAA 426
Db TAGACAAATTTGCTATCTCCCAAAATTTGCTTATCTACTCCCAAGACAACTCTGGAAGAA 5698

QY 427 TTCTGCAAAATCTTACGATTTGGGACAAACAAATGGCTACATCAAAATTTAGCACTTAA 486
Db TTCTGCAAAATCTTACGATTTGGGACAAACAAATGGCTACATCAAAATTTAGCACTTAA 5758

QY 487 TACAGACTCAGTTTGTGAGTGTATCATCCGATCCCATCCACCCAGTCACTCTCT 546
Db TACAGACTCAGTTTGTGAGTGTATCATCCGATCCCATCCACCCAGTCACTCTCT 5818

QY 547 TCTGGCCCAACCGAACCCCTCCATGCTTACCTGCTGAAACAAAGGACTTCTCTGTCA 606
Db TCTGGCCCAACCGAACCCCTCCATGCTTACCTGCTGAAACAAAGGACTTCTCTGTCA 5878

QY 607 ACCAAGACAAACAGGACAAAGCTGAAGAGTATTGAAGTTTCATGTATACAGATGCCA 666
Db ACCAAGACAAACAGGACAAAGCTGAAGAGTATTGAAGTTTCATGTATACAGATGCCA 5938

QY 667 AGGTTTATAGCTGTGATCTTCCCATCATCTGCTGGAAGACCTTGAAGTTTCAGAGC 726
Db AGGTTTATAGCTGTGATCTTCCCATCATCTGCTGGAAGACCTTGAAGTTTCAGAGC 5998

QY 727 TGAGTGTGAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTTCTGAATGCCAAACCT 786
Db TGAGTGTGAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTTCTGAATGCCAAACCT 6058

QY 787 GCAGAAATGTGACGGGAGCTCTTTTGAACCTGGGTGCTGCGCTTACCGCANTTACTGTC 846
Db GCAGAAATGTGACGGGAGCTCTTTTGAACCTGGGTGCTGCGCTTACCGCANTTACTGTC 6118

QY 847 TGCTGATTGATCCCAACCTGGGGGCGCTGTGACACACCTTTACTACTTTTCGATGCTCGG 906
Db TGCTGATTGATCCCAACCTGGGGGCGCTGTGACACACCTTTACTACTTTTCGATGCTCGG 6178

QY 907 GGGAGTGTGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG 966
Db GGGAGTGTGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG 6238

QY 967 TGAACAGAAAGTGTCTTACACCTGCTTCAAGAGGAACTTGAAGGCTCCCGGAGC 1026
Db TGAACAGAAAGTGTCTTACACCTGCTTCAAGAGGAACTTGAAGGCTCCCGGAGC 6298

QY 1027 GGTGACAGCTGCTGATACAGATCCCGAGCTGCTCAAGGGCTTACTTCCGGCGAGACTGTC 1086
Db GGTGACAGCTGCTGATACAGATCCCGAGCTGCTCAAGGGCTTACTTCCGGCGAGACTGTC 6358

QY 1087 AGGCTGCTGCTGGAGACCCAGATGCTGCTGTAATAACCGGGGTGTCTGCTTGTATCAT 1146
Db AGGCTGCTGCTGGAGACCCAGATGCTGCTGTAATAACCGGGGTGTCTGCTTGTATCAT 6418

QY 1147 ACTCGGCCACCGAGAGTGTAAATGCAACCGGCTTCAATGGGACGGGTGTGAGATGT 1206
Db ACTCGGCCACCGAGAGTGTAAATGCAACCGGCTTCAATGGGACGGGTGTGAGATGT 6478

QY 1207 GCTGCGCGGGAGATTGCGGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1266
Db GCTGCGCGGGAGATTGCGGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6538

QY 1267 GCGATGATGGCATCAGCGGCTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1326
Db GCGATGATGGCATCAGCGGCTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6598

QY 1327 CGTGTGACACTCAGGAGTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1386
Db CGTGTGACACTCAGGAGTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6658

QY 1387 CCTGTAAGAGAACCAACACGCTGTGATGTAACCTGGATTATGAAGTGCAGGAATCACAT 1446
Db CCTGTAAGAGAACCAACACGCTGTGATGTAACCTGGATTATGAAGTGCAGGAATCACAT 6718

QY 1447 GCACAGTTGTGATTTCTCAACAGGACAAACCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1506
Db GCACAGTTGTGATTTCTCAACAGGACAAACCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 6778

QY 1507 CCAGAAAGGCGACGAGGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1566
Db CCAGAAAGGCGACGAGGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6838

QY 1567 GCTGCAACAGATAGACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1626
Db GCTGCAACAGATAGACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6898

QY 1627 CCTGTAAGATGACAGGCGCGGCGGCAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686
Db CCTGTAAGATGACAGGCGCGGCGGCGGCAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686
```


Qy	67	TCGTGAAAGATCTGGTCCGCCAGGCCCTTCACTGTTTTTGGACCTTTTATCTCGACGCT	126
Db	5371	TCGTGAAAGATCTGGTCCGCCAGGCCCTTCACTGTTTTTGGACCTTTTATCTCGACGCT	5430
Qy	127	TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGTTTAAATGCCCGAGTTCTTC	186
Db	5431	TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGTTTAAATGCCCGAGTTCTTC	5490
Qy	187	GGTACCATGTGTCGCTGCCACAGCTGCTTCTGGAAACCTGAAATTCATCTCAAAATG	246
Db	5491	GGTACCATGTGTCGCTGCCACAGCTGCTTCTGGAAACCTGAAATTCATCTCAAAATG	5550
Qy	247	CTACTTCCCTCCAAAGGAGCAATAGTCACTCCCGTCTCTCAGAGCACGGTGATATAA	306
Db	5551	CTACTTCCCTCCAAAGGAGCAATAGTCACTCCCGTCTCTCAGAGCACGGTGATATAA	5610
Qy	307	ATAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATGTTCAATATCA	366
Db	5611	ACAATAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATGTTCAATATCA	5670
Qy	367	TAGACAAATTCGCTATCTCCAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGAA	426
Db	5671	TAGACAAATTCGCTATCTCCAAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGAA	5730
Qy	427	TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAAATTTAGCAACTTAA	486
Db	5731	TTCTGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAAATTTAGCAACTTAA	5790
Qy	487	TACAGGACTCAGGTTTGCTGAGTGTCATCACCAGTCCATCCACACCCAGTCACTCTCT	546
Db	5791	TACAGGACTCAGGTTTGCTGAGTGTCATCACCAGTCCATCCACACCCAGTCACTCTCT	5850
Qy	547	TCTGSCCCACAGCAAA-----GCCCTCCATGCCCCCTACTCTGCTGAAACAAAGGACT	597
Db	5851	TCTGSCCCACAGCAAA-----GCCCTCCATGCCCCCTACTCTGCTGAAACAAAGGACT	5910
Qy	598	TCCTGTTCAACCAAGACAAAGGAGTAAAGGAGTATTTGAAAGTTTCATGTGATAC	657
Db	5911	TCCTGTTCAACCAAGACAAAGGAGTAAAGGAGTATTTGAAAGTTTCATGTGATAC	5970
Qy	658	GAGATGCCAAGTTTACTGTGATCTTCCACATCCACTGCTGCGGAGACCCCTGCAAG	717
Db	5971	GAGATGCCAAGTTTACTGTGATCTTCCACATCCACTGCTGCGGAGACCCCTGCAAG	6030
Qy	718	GTTTCAGAGCTGAGTGAATTTGAGCTGGCAGGACATCGGTGACCTTTCTGTAATG	777
Db	6031	GTTTCAGAGCTGAGTGAATTTGAGCTGGCAGGACATCGGTGACCTTTCTGTAATG	6090
Qy	778	GCCAAACCTGCAGAAATGTGCAGCGGGAGCTTTGTTTGACCTGGGTGGCCATCGGCA	837
Db	6091	GCCAAACCTGCAGAAATGTGCAGCGGGAGCTTTGTTTGACCTGGGTGGCCATCGGCA	6150
Qy	838	TTGACTGTCTGATGATTCACACCTGGGGGGCGCTGTGACACCTTTACTATTTCG	897
Db	6151	TTGACTGTCTGATGATTCACACCTGGGGGGCGCTGTGACACCTTTACTATTTCG	6210
Qy	898	ATGCGCTCGGGGAGTGGAGCTCTGTCAATCTCCAGCTGCCAAAGTGGAGTAAAC	957
Db	6211	ATGCGCTCGGGGAGTGGAGCTCTGTCAATCTCCAGCTGCCAAAGTGGAGTAAAC	6270
Qy	958	CAAAGGTTGTAAGCAGAGTGTCTCTAACACCTGCCCTTCAAGAGGAACCTGGAGGCT	1017
Db	6271	CAAAGGTTGTAAGCAGAGTGTCTCTAACACCTGCCCTTCAAGAGGAACCTGGAGGCT	6330
Qy	1018	GCCGGGAGCGGTGACCTGGTGATACAGATCCCGAGTGTGCAAGGGCTACTTTCGGGC	1077
Db	6331	GCCGGGAGCGGTGACCTGGTGATACAGATCCCGAGTGTGCAAGGGCTACTTTCGGGC	6390
Qy	1078	GAGACTGTCAAGGCTGCCCTGGAGGACAGATGCCCGCTGTAAATAACCGGGGTGTCTGCC	1137
Db	6391	GAGACTGTCAAGGCTGCCCTGGAGGACAGATGCCCGCTGTAAATAACCGGGGTGTCTGCC	6450
Qy	1138	TTGATCAGTACTCGGCCACCGGAGAGTGAATGCAACACCGGCTTCAATGGGACGGGCT	1197

7424	GCTCAGCTCGAGCGCGTGCAATTTCTAGAACCTGACTGACCTGTGCATCGCGGCACCC	7483
2278	TCCTTTGTGCCACAGAACAGTGGGCTGGGGAGATGAGACTTGTCTCTGGCGGGGACATCG	2337
7484	TCCTTTGTGCCACAGAACAGTGGGCTGGGGAGATGAGACTTGTCTCTGGCGGGGACATCG	7543
2338	AGCACACCTCGCCAAATGTCTAGCANGTGTTTCTCAATGACCTTGTCTCAATGGCACCAACC	2397
7544	AGCACACCTCGCCAAATGTCTAGCANGTGTTTCTCAATGACCTTGTCTCAATGGCACCAACC	7603
2398	TGCAAAACGAGGCTGGGAAGCAAGCTGCTCATCACTGTCGACGACGACCACTCCACCCGA	2457
7604	TGCAAAACGAGGCTGGGAAGCAAGCTGCTCATCACTGTCGACGACGACCACTCCACCCGA	7663
2458	CGGAGACCAAGTGTGTTGATGGAAAGCCATTTCTGCAGTGGGAGCATTTTGGCCTCCAAATG	2517
7664	CGGAGACCAAGTGTGTTGATGGAAAGCCATTTCTGCAGTGGGAGCATTTTGGCCTCCAAATG	7723
2518	GGATCATTCATGTCATTTCCAGGCTTTAAAGCACCCCTGCCCCCGTGACCTTGACCC	2577
7724	GGATCATTCATGTCATTTCCAGGCTTTAAAGCACCCCTGCCCCCGTGACCTTGACCC	7783
2578	ACACTGGCTTGGGACGAGGATCTTCTTGTGCAATCTCTGTGTGACCTTGGGCTGTGACCT	2637
7784	ACACTGGCTTGGGACGAGGATCTTCTTGTGCAATCTCTGTGTGACCTTGGGCTGTGACCT	7843
2638	TGGCTGCTTACTCTACTTTGGATAAACCGGAGAACAAATCGGCTTCCAGCATTTTGAGT	2697
7844	TGGCTGCTTACTCTACTTTGGATAAACCGGAGAACAAATCGGCTTCCAGCATTTTGAGT	7903
2698	CGGAAGAGGACATTAAGTTGTGAGCTCTTGGCAGCAGCAGCTGAGAAATCTCGAAC	2757
7904	CGGAAGAGGACATTAAGTTGTGAGCTCTTGGCAGCAGCAGCTGAGAAATCTCGAAC	7963
2758	CCTTGTATGAGAGCAACCTCAGCTCCCCAGAACCTTCTACGACCCCTTCAACGACT	2817
7964	CCTTGTATGAGAGCAACCTCAGCTCCCCAGAACCTTCTACGACCCCTTCAACGACT	8023
2818	CTGAGAACCGCAGCTTGGGGCAATGACCCCTTGAGGACACTGTGAGGCGCTGGACGG	2877
8024	CTGAGAACCGCAGCTTGGGGCAATGACCCCTTGAGGACACTGTGAGGCGCTGGACGG	8083
2878	AGATGCCAGCCATCACTCACTGCGACCTGGGCCATCAACTGTCAATCTTCAGCACCACTT	2937
8084	AGATGCCAGCCATCACTCACTGCGACCTGGGCCATCAACTGTCAATCTTCAGCACCACTT	8143
2938	GCCTTTTAGGAAACGTAAGTCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG	2997
8144	GCCTTTTAGGAAACGTAAGTCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG	8203
2998	ATCTGGGGGTGTTCTGTGGGTGAGAGATGTCTGTGCTGCCACCCAGTACAGCTTCC	3057
8204	ATCTGGGGGTGTTCTGTGGGTGAGAGATGTCTGTGCTGCCACCCAGTACAGCTTCC	8263
3058	TCCTCTGACCTTTGGCTCTTCTTCTCTTTGTACTCTTCAGCTGGCACCTGCTCCATCTG	3117
8264	TCCTCTGACCTTTGGCTCTTCTTCTCTTTGTACTCTTCAGCTGGCACCTGCTCCATCTG	8323
3118	CCCTACATGATGGTAACCTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCTCTTTG	3177
8324	CCCTACATGATGGTAACCTGTGATCTTTCTTCCCTGTTAGATTGTAAGCCTCCTCTTTG	8383
3178	TATCCAGGCCCTTAGCCAGTGTCTGACACAGGAACCTGTGCAATAAAGGTTTATGGAA	3237
8384	TATCCAGGCCCTTAGCCAGTGTCTGACACAGGAACCTGTGCAATAAAGGTTTATGGAA	8443
3238	CAGAAAAAATAAAAAAAAAAAAAA	3260
8444	CAGAAAAAAGTCAAAAAAAAAAAAA	8466

LOCUS	AX149461	3625 bp	DNA	linear	PAT 08-JUN-2001
DEFINITION	Sequence 17 from Patent WO0136638.				
ACCESSION	AX149461				
VERSION	AX149461.1 GI:14347936				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 Shimkets,R.A., Lichenstein,H., Vernet,C. and Fernandes,E.				
TITLE	Polypeptides and nucleic acids encoding same				
JOURNAL	Patent: WO 0136638-A 17 25-MAY-2001;				
FEATURES	Curagen Corporation (US)				
source	Location/Qualifiers				
	1..3625				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
	Query Match 86.8%; Score 2828.8; DB 6; Length 3625;				
	Best Local Similarity 95.1%; Pred. No. 0;				
	Matches 2985; Conservative 0; Mismatches 43; Indels 111; Gaps 2;				
Qy	122 AGCCTTTGATGAGGAAGCTCGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGGT 181				
Db	569 ACCCTCTGATCTTCTGTACAGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGGT 628				
Qy	182 TCITTCGGTACCATGTGGTCGCTGCCACCAAGCTGCTTCTGGAACCTGAAATGATCTC 241				
Db	629 TCITTCGGTACCATGTGGTCGCTGCCACCAAGCTGCTTCTGGAACCTGAAATGATCTC 688				
Qy	242 AAATGCTACTTCCCTCCAAAGAGAGCAATAGTACATCTCCGTCCTCTCAGAGCAGCGTGA 301				
Db	689 AAATGCTACTTCCCTCCAAAGAGAGCAATAGTACATCTCCGTCCTCTCAGAGCAGCGTGA 748				
Qy	302 TATAAATAAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTTGTCA 361				
Db	749 TATAAATAAAGGCTAAGATCATATCCAGTGATATCATCAGTACTAATGGGATTTGTCA 808				
Qy	362 TATCATAGACAAATTGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGG 421				
Db	809 TATCATAGACAAATTGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGG 868				
Qy	422 AAGAAATTCGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAAATTAGCAA 481				
Db	869 AAGAAATTCGCAAAATCTTACGACTTTGGCAACAAACAATGGCTACATCAAAATTAGCAA 928				
Qy	482 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCCGATCCCATCCACACCCAGTCAC 541				
Db	929 CTTAATACAGGACTCAGGTTTGCTGAGTGTCATCCGATCCCATCCACACCCAGTCAC 988				
Qy	542 TCTCTTCTGGCCCAACCGAACCCCTCCATGCCCTACCTGCTGCTGCTGCTGCTGCTGCT 601				
Db	989 TCTCTTCTGGCCCAACCGAACCCCTCCATGCCCTACCTGCTGCTGCTGCTGCTGCTGCT 1048				
Qy	602 GTTCAACCAAGACAAACAAGGACAAAGCTGGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 661				
Db	1049 GTTCAACCAAGACAAACAAGGACAAAGCTGGAAGGAGTATTTGAAGTTTCATGTGATACGAGA 1108				
Qy	662 TGCACAGGTTTACGCTGTGGATCTTCCACATCCACTGCTGGAAGACCTCGAAGGTTTC 721				
Db	1109 TGCACAGGTTTACGCTGTGGATCTTCCACATCCACTGCTGGAAGACCTCGAAGGTTTC 1168				
Qy	722 AGAGCTGAGTGTGAAATGGAGCTGGCGGGACATCGGTGACCTCTTTCTGAATGGCCA 781				
Db	1169 AGAGCTGAGTGTGAAATGGAGCTGGCGGGACATCGGTGACCTCTTTCTGAATGGCCA 1228				
Qy	782 AACCTCGAATTTGTGACGGGAGCTCTTGTGTTGACCTGGGTGTGGCCCTACGGCAATTGA 841				
Db	1229 AACCTCGAATTTGTGACGGGAGCTCTTGTGTTGACCTGGGTGTGGCCCTACGGCAATTGA 1288				

Qy	842	CTGTCTGCTGATTGATCCCACTCGGGGGCCGTGTGACACTTTTACTACTTTTGATGC	901
Db	1289	CTGTCTGCTGATTGATCCCACTCGGGGGCCGTGTGACACTTTTACTACTTTTGATGC	1348
Qy	902	CTCGGGGGAGTGTGGGAGCTGTGTCAATATCTCCAGCTGCCAAGGTGAGTGAACCAAA	961
Db	1349	CTCGGGGGAGTGTGGGAGCTGTGTCAATATCTCCAGCTGCCAAGGTGAGTGAACCAAA	1408
Qy	962	GGGTGTCAAGCAGAGTGTCTTACACCTGCCCTTCAAGAGAACTCTGAAGGCTGCCG	1021
Db	1409	GGGTGTGAAGCAGAGTGTCTTACACCTGCCCTTCAAGAGAACTCTGAAGGCTGCCG	1468
Qy	1022	GGAGCGTGCAGCCTCGTGTGATACAGATPCCCAAGTGTGCAAGGGCTACTTCGGGCGAGA	1081
Db	1469	GGAGCGTGCAGCCTCGTGTGATACAGATPCCCAAGTGTGCAAGGGCTACTTCGGGCGAGA	1528
Qy	1082	CTGTTCAGGCTGCCCTGGAGGAACAGATGCCCGTGTATATACCGGGGTGTCTGCCCTTGA	1141
Db	1529	CTGTTCAGGCTGCCCTGGAGGAACAGATGCCCGTGTATATATACCGGGGTGTCTGCCCTTGA	1584
Qy	1142	TCAGTACTTCGGCCACCGGAGAGTGAATCAACACCGGCTTCAATGGAGCGCGTGTGA	1201
Db	1585	TCAGTACTTCGGCCACCGGAGAGTGAATCAACACCGGCTTCAATGGAGCGCGTGTGA	1644
Qy	1202	GATGTGCTGGCGGGGAGATTCGGGCGCTGATTGTCTGCCCTGTGGTGTCTAGACACCG	1261
Db	1645	GATGTGCTGGCGGGGAGATTCGGGCGCTGATTGTCTGCCCTGTGGTGTCTAGACACCG	1704
Qy	1262	ACAGTGGGATGATGGCATCACCGGCTCGGGGAGTGCCTCTGTGTAAACCGGGTGGACGG	1321
Db	1705	ACAGTGGGATGATGGCATCACCGGCTCGGGGAGTGCCTCTGTGTAAACCGGGTGGACGG	1764
Qy	1322	CCCCTGTGTACACTCAGGCAGTTTTCGCTGCAGTGTGTACGCTCTTGTGTCTCTCA	1381
Db	1765	CCCCTGTGTACACTCAGGCAGTTTTCGCTGCAGTGTGTACGCTCTTGTGTCTCTCA	1824
Qy	1382	TGCCACCTGTGAAGGAGAACACACGTGTGAGTGTAACTCGATATATGAAGGTGACCGAAT	1441
Db	1825	TGCCACCTGTGAAGGAGAACACACGTGTGAGTGTAACTCGATATATGAAGGTGACCGAAT	1884
Qy	1442	CACATGCACAGTTGTGATTTCTGCAACACGAGCAACCGGGGCTGTGCAAGGTGGCCAG	1501
Db	1885	CACATGCACAGTTGTGATTTCTGCAACACGAGCAACCGGGGCTGTGCAAGGTGGCCAG	1944
Qy	1502	ATGCTCCCAGAAGGGCACGAGGTCTCTCTGCAGTGCACAGAGGATACAAAGGGACACGG	1561
Db	1945	ATGCTCCCAGAAGGGCACGAGGTCTCTCTGCAGTGCACAGAGGATACAAAGGGACACGG	2004
Qy	1562	GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCTTTAACGGAGGTGTCCACGACA	1621
Db	2005	GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCTTTAACGGAGGTGTCCACGACA	2064
Qy	1622	CGCACCTGTGAAGATGACAGGCCCGGGCAGACACAGTGTGAGTGTAAAGTCACTATGT	1681
Db	2065	CGCACCTGTGAAGATGACAGGCCCGGGCAGACACAGTGTGAGTGTAAAGTCACTATGT	2124
Qy	1682	CGGAGATGGCTGAACCTGTGAGCCGGAGCAGCTCCCACTTACCCGCTGTCTTACAGACAA	1741
Db	2125	CGGAGATGGCTGAACCTGTGAGCCGGAGCAGCTCCCACTTACCCGCTGTCTTACAGACAA	2184
Qy	1742	TGGGCAGTGCATGCAGACGGCAAAATGTGTGCAGCTCCACTTCAGGATACCACTGTGG	1801
Db	2185	TGGGCAGTGCATGCAGACGGCAAAATGTGTGCAGCTCCACTTCAGGATACCACTGTGG	2244
Qy	1802	GGTGTTCATCTACGCTCCCACTGGGCGAGTATAAGCTGACCTTTGACAAAGCCAGAGA	1861
Db	2245	GGTGTTCATCTACGCTCCCACTGGGCGAGTATAAGCTGACCTTTGACAAAGCCAGAGA	2304
Qy	1862	GGCTGTGCGCAACGAAGCTGGCAACCATGGCAACCTTACAACGAGCTCTCTTATGCCAGAA	1921
Db	2305	GGCTGTGCGCAACGAAGCTGGCAACCATGGCAACCTTACAACGAGCTCTCTTATGCCAGAA	2364
Qy	1922	GGCCCAAGTACCACTGTGTCTCAGCAGGTGTGCTGGAGACCGGGCGGGTGTGCTTACCCAC	1981

Db	2365	GGCCAAAGTACCACTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCATTACCCAC	2424
QY	1982	AGCCTTCGGCTCCAGAACTGTGGCTCTGGTGTGGATAGTGGACTATGGACCTAG	2041
Db	2425	AGCCTTCGGCTCCAGAACTGTGGCTCTGGTGTGGATAGTGGACTATGGACCTAG	2484
QY	2042	ACCCAAACAAGTGAATGTGGATGTCTTCTGTATCGGATGAAGATGTGAACATGCAC	2101
Db	2485	ACCCAAACAAGTGAATGTGGATGTCTTCTGTATCGGATGAAA-----	2530
QY	2102	CTGCAAGGTGGGCTATGTGGGAGATGGCTTCTCATGCACTGGGAACCTGCTCAGGTCCCT	2161
Db	2531	-----	2530
QY	2162	GATGTCTTCCCCTCACTCACAACTTCCTGACGGAAGTGTGGCCTATTCCAAACAGCTTC	2221
Db	2531	-----GGAAAGTGTGGCCTATTCCAAACAGCTTC	2557
QY	2222	AGCTCAGGCGGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGACACCTCTTT	2281
Db	2558	AGCTCAGGCGGTGCATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGACACCTCTTT	2617
QY	2282	TGTGCCACAAACAGTGGGCTGGGGAGAAATGAGACCTTTGTCTGGCGGGACATCGAGCA	2341
Db	2618	TGTGCCACAAACAGTGGGCTGGGGAGAAATGAGACCTTTGTCTGGCGGGACATCGAGCA	2677
QY	2342	CGACCTGCGCAATGTGAGCAATGTTTTTCTACAATGACCTTGTCAATGGCACCAACCTGCA	2401
Db	2678	CCACCTGCGCAATGTGAGCAATGTTTTTCTACAATGACCTTGTCAATGGCACCAACCTGCA	2737
QY	2402	AACGAGGCTGGGAAGCAAGCTGCTCATCTGCGCAGCAGGACCCACTCCAAACCGACGGA	2461
Db	2738	AACGAGGCTGGGAAGCAAGCTGCTCATCTGCGCAGCAGGACCCACTCCAAACCGACGGA	2797
QY	2462	GACAGGTTTGTGTGATGGAAGAGCAATTCGCACTGGGACATCTTTTGCTTCCATGGGAT	2521
Db	2798	GACAGGTTTGTGTGATGGAAGAGCAATTCGCACTGGGACATCTTTTGCTTCCATGGGAT	2857
QY	2522	CAATCATGTCAATTTCCAGGCGCTTTAAAGCACACCCCTGCCCGGTGACCTTGACCCACAC	2581
Db	2858	CAATCATGTCAATTTCCAGGCGCTTTAAAGCACACCCCTGCCCGGTGACCTTGACCCACAC	2917
QY	2582	TGGCTTGGGAGCAGGATCTTTTGCATCATCTTGGTGACTGGGGCTGTTCGCTTGGC	2641
Db	2918	TGGCTTGGGAGCAGGATCTTTTGCATCATCTTGGTGACTGGGGCTGTTCGCTTGGC	2977
QY	2642	TGCTTACTCTACTTTTCGGATTAACCGGAGAAACAATGGGCTCCAGACATTTTCAGTCGGA	2701
Db	2978	TGCTTACTCTACTTTTCGGATTAACCGGAGAAACAATGGGCTCCAGACATTTTCAGTCGGA	3037
QY	2702	AGAGGACATTAATGTTTCAGCTCTTGGAAGCAGCAGCCTGAGAATATCTCGAACCCCTTT	2761
Db	3038	AGAGGACATTAATGTTTCAGCTCTTGGAAGCAGCAGCCTGAGAATATCTCGAACCCCTTT	3097
QY	2762	GTATGAGAGCAACAACCTCAGCTCCCCCAGAACCTTCTCAGACCCCTTTCAGGACTCTGA	2821
Db	3098	GTATGAGAGCAACAACCTCAGCTCCCCCAGAACCTTCTCAGACCCCTTTCAGGACTCTGA	3157
QY	2822	AGACGCGACGTTGAGGGCAATGACCCCTTGAGGACACTCTGAGGGCTCGGACGGGAGAT	2881
Db	3158	AGAACGCGAGCTTGAGGGCAATGACCCCTTGAGGACACTCTGAGGGCTCGGACGGGAGAT	3217
QY	2882	GCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAAATTTCTCAGACCAAGTTGCCT	2941
Db	3218	GCCAGCCATCACTCACTGCCACCTGGGCCATCAACTGTGAAATTTCTCAGACCAAGTTGCCT	3277
QY	2942	TTTATGGAACGTAAGTCTTTTATGACCTCAGAGGCCATACCTCATCTCTCTGGCTGATCT	3001
Db	3278	TTTATGGAACGTAAGTCTTTTATGACCTCAGAGGCCATACCTCATCTCTCTGGCTGATCT	3337
QY	3002	GGGGGTTGTTTTCTGTGGGTGAGAGATGTGTTGTGTGCCCAACCCAGTACAGCTTTCCTCCT	3061

QY	431	GCATAATCTTACGCTTTGGCAACAAACAATGGCTACATCAAAATTTAGCAACTTAATACA	490	Db	6385	CCAGAGGGCACACAAGTCTCTCTGAGCTGCAGAGGGGTACAGGGGGATGGCCACAG	6444
Db	5305	GCTAAACCTTACTACAGTGGCAGCAAAACCATGGATATACCAAAATTCAGAAATTTGATACA	5364	QY	1568	CTGCACAGAGATAGACCCCTGTGCAGACGGCTTTAACGAGGGTGTCTACGAGCAGCCAC	1627
QY	491	GGACTCAGGTTTGTCTGAGTGTCAATCACCGATCCCAATCCACACCCCAAGTCACTCTCTCTG	550	Db	6445	CTGCACAGAGATAGACCCCTGTGCAGACGGTGTCAATGGTGGATGTCTATGACATGCTTAC	6504
Db	5365	GGACTCAGGTTTGTCTGAGTGTCAATCACCGATCCCAATCCACACCCCAAGTCACTCTCTCTG	5424	QY	1628	CTGTAAGATGACAGGCCCGGCAAGCAACAAGTGTGAGTGTAAAGTCACTATGTCTGGAGA	1687
QY	551	GCCACCGACCAAGCCCTCCATGCTCTACCTGTGAAACAAAGAGACTTCTGTTCACACA	610	Db	6505	CTGCAGATGACAGGCCCGGCAAGCAACAAGTGTGAGTGTAAAGTCACTATGTCTGGAGA	6564
Db	5425	GCTACCGACCAAGCCCTCCAGGCTTGCCTCAGGAGCAGCAGGACTTCTGTTCATATGA	5484	QY	1688	TGGGCTGAATGTGAGCCCGGAGCTGCCATTCACCGCTGTCTTACAGGACAATGGGCA	1747
QY	611	AGACAACAGCAGCAGCTGAGGAGTATTTGAAGTTTCATGTGATACGAGATGCAAGT	670	Db	6565	TGGACCGGACTGTGAGCCCGGCAAGCAACAAGTGTGAGTGTAAAGTCACTATGTCTGGAGA	6624
Db	5485	AGACAACAGCAGCAGCTGAGGAGTATTTGAAGTTTCATGTGATACGAGATGCAAGT	5544	QY	1748	GTGCCATGACAGCCCAATGTGTGAGCTCCACCTTCAGGATACACCTGTTGGGGTGT	1807
QY	671	TTTAGCTGTGATCTTCCACATCCACTGCTCTGGAAGACCTTCGAAGSTTCAGAGCTGAG	730	Db	6625	ATGCCACCCAGATGCAAACTGTGTGAGCTTCAATTCAGGACACTTCTGTTGGGGTAT	6684
Db	5545	TTTAGCTTCAGACCTCCCGAGTCTCTCTCTGGAAGACCTTCGAAGSTTCAGAGCTGAG	5604	QY	1808	CCATCTAGCTCCCACTGGGCGAGTATAAGCTGACCTTTTGACAAAGCCAGAGAGCCCTG	1867
QY	731	TGTGAATGTGAGCTGGCAGGAGCATCGGTGACCTTCTGAATGGCCAAACCTGCAG	790	Db	6685	CCACCTAGCTCCCACTGGGCGAGTATAAGCTGACCTTTTGACAAAGCCAGAGAGCCCTG	6744
Db	5605	TGTGAGGTGTGAACTGGCAGTGTGTCGGAGAGCTTCTTCAATGGACAAATGTGCAG	5664	QY	1868	TGCCAACGAGCTGGGACCATGGGAAACCTACAAACAGCTCTCTATGCCCCAGAGCCAA	1927
QY	791	AATTGTGACGCGGAGCTCTGTTTGAACCTGTGGCTACGGCAATTGACTGTCTGCT	850	Db	6745	TGCTAAGGAAGCTGCATGCATAGCCACCTTACAAACAGCTCTCTATGCCCCAGAGCCAA	6804
Db	5665	GATTATACAGCGGAGACTCTGTTTGAACGAGGTGTGGCTATGSCATTTGACTGCTGCT	5724	QY	1928	GTACCACTGTGCTCAGCAGGCTGGCTGGAGACCGGGGGTGGCTACCCACAGCCCT	1987
QY	851	GATTGATCCACCTCGGGGGCGCTGTGACACCTTTACTACTTTGATGCTCCGGGGA	910	Db	6805	GTATCACCTCTGCTCAGCTGGCTGGAGAGTGGGGGGTGGCTACCCACAGTCTA	5864
Db	5725	AATGATCTACCGAAGGTGGCGATGTGACACTTTTACTACTTCAATATTCGGGGGA	5784	QY	1988	CGCCTCCACAGAACTGTGGCTCTGGTGTGGGATAGTGGACTATGGACCTTAGACCCAA	2047
QY	911	GTGTGGAGCTGTGCAATCTCCAGCTGCCCCAGGTGGATGAGTAACCAAGGTTGAA	970	Db	6865	CGCCTCTAAGAAATGTG---CAAACTTGTGGGATGTAGACTACGGAACAGGACCAA	6921
Db	5785	ATGTGGAGCTGTTTCTCACTCCAGATGCCACTACAGAGCAAAACCAAGGGTGTGAG	5844	QY	2048	CAAGAGTGAATGTGGGATGTCTTCTGTATCGGATGAAGATGGAATGCACTGACCTGCA	2107
QY	971	GCAGAGTGTCTTACAACC---TGCCCTTCAAGAGGAACCTTGAAGGCTGCCGGAGCG	1027	Db	6922	CAAGAGTGAATGTGGGAGCTCTTCTGTGTACAGAAATGAAGATGTTAACTGACCTGCA	6981
Db	5845	GAAAGATGTATCTACAAACCGTTACTTTCAGAGGATGTGGAAGGCTGTCAAGACCT	5904	QY	2108	GGTGGCTATGTGGGAGATGGCTTCTCATGAGTGGGAACCTGCTGCAGTCTCTGATGTC	2167
QY	1028	GTGACCTGTGTGATACAGATCCCAAGTGTGTGCAAGGCTACTTCGGCGAGACTGTCA	1087	Db	6982	GGAGGCTACGTGGGGGATGGCTTCTCTGCAATGGGAACCTGCTGCAAGTCTCTATGTC	7041
Db	5905	GTGACCTGTGTGATACAGATCCCAAGTGTGTGCAAGGCTACTTCGGCGAGACTGTCA	5964	QY	2168	CTTCCCTCACTCACAACTTCTGACGGAAGTGTGGCTATTCCAAAGCTCAGCTCG	2227
QY	1088	GGCTGCTGTGAGGACAGATGCCCGGTGTATACCGGGTGTCTGCTTGTATCAGTA	1147	Db	7042	CTTCCCTCACTCACAACTTCTGACGAGGCTGTGCTTTTTCAGGAGCTCAGCCCA	7101
Db	5965	GGCTGCTGTGAGGACAGACACACCGTGTAAACCGGGGATGTGTTATGATCAGTA	6024	QY	2228	AGGCTGTGATTTTAGAACACTGACTGACCTGTCCATCCGGCGACCCCTCTTTGTGCC	2287
QY	1148	CTCGGCACCGGAGTGTAAATGCAACACCGGCTTCAATGGGACCGCGTGTGAGATGTG	1207	Db	7102	AGGCGGGGGTTTTTGAACACCTGACTGACCTGTCCATCAGTGGCACCTGTTGTGCC	7161
Db	6025	CAAAACCAAGACAGTGCATATGCTACTGTTTCAAGGACAGCTGTGAGCTCTG	6084	QY	2288	ACAGAACAGTGGGCGAGAAATGAGACCTTGTCTGGCGGGGACATCGAGCACCCT	2347
QY	1208	CTGGCGGGGAGATTCGGGCTGAATGTGTGCTGTGCTGTGCTCAGACCGGACAGTG	1267	Db	7162	ACAGAACAGTGGGCTACCCAAAATAGAGCTGTCTGGCGGGACATTTGAGCACCCT	7221
Db	6085	CTGCTGGGAGATTTGGTCTGACTGTCAACCTGTGGCTGTCTGAGCATGGGAGTG	6144	QY	2348	CGCAATGTGAGATGTTTTTCTACAATGACCTGTCAATGGCACCACTTGAACAGAG	2407
QY	1268	CGATGTGATCATCAGGGCTCCGGGAGTGCCTCTGTGAACCGGGTGTGACAGCCCTC	1327	Db	7222	CACATATGTCAGCTCTCTCTTTTACGATGACCTTGTCAATGTGTACGCTCTGAAAGTAG	7281
Db	6145	TGACGAGGGGATCACAGGCTCCGGGAGTGCCTCTGTGAAGAGGGTGTGACAGCCCT	6204	QY	2408	GGTGGGAAGCAGCTGCTCATCTACCTCCAGCAGGACCCACCTCCACCGAGGAGCCAG	2467
QY	1328	GTGTGACACTCAGGAGTTTGGCTGTCAGTGTGTAGCCCTCTCTGTTGCTCATGCCAC	1387	Db	7282	GCTGGGAAGCCAACTGCTCATCTCCAGCAGGACCCAGCTCCAC---AAGAGGCCAG	7338
Db	6205	CTGTGATGCTCCACAGTTGTGATTCAGTGTGCATACCTGCTTGTTCATGACGCAAC	6264	QY	2468	GTGTTGTGATGGAAGAGCCATCTGCAGTGGGACATCTTTTGGCTCCAAATGGGATCA	2527
QY	1388	CTGTAAAGGAGAAACACAGTGTGAGTGTAAACCTGGATTAAGAGGTGACGGAATCATG	1447	Db	7339	GTTCGTGATGGAAGAGCCATCTGCAGTGGGACATCAATGCTCTAAACGGGGTCTCCA	7398
Db	6265	CTGTATGGAGAAACACACATGTGTGTACTTTGAATTAAGAGGTGACGGAATCATG	6324	QY	2528	TGTCATTTCCAGCCCTTTAAAGCAACCCCTGCCCCCGTGAACCTTGACCCACACTGGCT	2587
QY	1448	CACAGTGTGATTTCTGAAAACAGGACACCGGGCTGTGAAAAGTGGCCAGATGCTC	1507	Db	7399	TATCAUTTTCTGAACCTTTGAAAGCTCTCCCAACCGGCGCAACCGGCTGCCACTCTGGCCT	7458
Db	6325	CACAGTGTGATTTCTGAAAACAGAAACATGGGGCTGTGGAAGTGTGCCAAGTGTCTC	6384	QY	2588	GGAGAGGGGATCTTCTTTGCCATCATCTGTGTGACTGTGGGGCTGTGCTTGGCTGTCTA	2647
QY	1508	CCAGAGGGGACGAAGTCTCTCTGACGTGCCAGAGGGGATACAAAGGGGACGGGCACAG	1567				


```
Db 7459 AGGAACAGGCAATTTCTGGCTGTCTGCTCTGTTACTGTGCGATTGCTCTGGCTGCCTA 7518
QY 2648 CTCCTACTTTCCGATAAACCAGAGAAACAATCGGCTTCCAGCATTTTGTAGTCGGAAGAGGA 2707
Db 7519 CTCCTACTTCCGGCTAAACACAGAGAAACAATCGGCTTCCGGCGTTTGTAGTCAGAAAGATGA 7578
QY 2708 CATTATATGTTGAGCTCTTGGCAACAGAGAGCTGAGATATCTCGAACCCCTTGTATGA 2767
Db 7579 CATTGATGCTTGGGCTTTGGCAAGCAGAGCTGAGATATCAAAACCCCTTGTATGA 7638
QY 2768 GAGCAAAACCTCAGCTCCGCCAGAAACCTTCCTACGACCCCTTCACGAGCTCTGAAGAAG 2827
Db 7639 GACCTCAAGCGCGGAGCCCCGAGCCCTCTCTGTGACCCCTTCACGACTCTGGAGAAG 7698
QY 2828 GAGCTTGAAGGCAATGACCCCTTGAAGACACTGTGAGGCGCTTGAAGGAGATCCAGC 2887
Db 7699 GAGCTGAGGAACAGTACCCCTCTTGGGGCACTG---AGGTCCTGATGTGAGAAGCCAGC 7755
QY 2888 CATCACTCACTGCCACCTGGGCCATCA-----ACTGTGAATCTCAGCACC 2933
Db 7756 CAACAGCCACAGCCAGCTGAGCCCTGAGCCATCACGGTTACACGGTGACTCCCTGCTCC 7815
QY 2934 AGTTGCTCTTTAG---GAACCTAAAGTCTTTAAGCACTCAGAACCCATACCTCATCTCT 2990
Db 7816 AGCTGCTACTTGGATCATTTGTTTAAAGTATGATAAATCAGAACCCATACCTCACCCCT 7875
QY 2991 CTGGCTGATCTGGGGGTTTCTCTGGGTGAGAGATGTGTGCTGTGTCGCCACCCAGTAC 3050
Db 7876 CTGGTTAGCTTGGGGTTTCTGCTCATGGGTGAGGGGACATGTTTCCAGGACACCACTATAC 7935
QY 3051 AGCTTCTCTCTGACCCCTTGGCTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3110
Db 7936 CTGCTCTCTCTTGGGCTCTCACTGTGGCT--CTCTGACCTTCCAGATGGTGTGGCC 7993
QY 3111 CATTCTGCCCTACATGATGGTAACTGTGATCTT 3144
Db 7994 TGTCTGCCCTATCTG-TGGGTAGCTGTGACCCT 8026
```

```
RESULT 14
AY007370 4706 bp mRNA linear ROD 23-OCT-2002
LOCUS Rattus norvegicus hyaluronan receptor for endocytosis HARE.
DEFINITION precursor, mRNA, partial cds.
ACCESSION AY007370
VERSION AY007370.1 GI:24285892
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 4706)
AUTHORS Zhou, B., Weigel, J. A., Saxena, A. and Weigel, P. H.
TITLE Molecular Cloning and Functional Expression of the Rat 175-kDa
Hyaluronan Receptor for Endocytosis
JOURNAL Mol. Biol. Cell 13 (8), 2853-2868 (2002)
MEDLINE 22169209
PUBMED 12181351
REFERENCE 2 (bases 1 to 4706)
AUTHORS Zhou, B., Weigel, J. A. and Weigel, P. H.
TITLE Direct Submission
JOURNAL Submitted (25-AUG-2000) Department of Biochemistry and Molecular
Biology, University of Oklahoma Health Sciences Center, 940 Stanton
L. Young, BMSB 860, Oklahoma City, OK 73104, USA
Location/Qualifiers
FEATURES
source
1. 4706
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
<1..4296
/codon_start=1
CDS
1875 GGACTCAGGCTTGTCTGTCAGTCACTGACTCCATCCACACCCAGTCACTGTCTTCTG 1934
```

```
/product="hyaluronan receptor for endocytosis HARE
precursor"
/protein_id="AAG13634.1"
/db_xref="GI:24285893"
/translation="SLPSLLTRLEQMPDYSIFRGVIIHYNLASATESADAYTVFVFN
EATENYIREKATSLKEDILRYHVLGKLEKLNLDHNGHRETMGLFSGYLLAFRLND
QLYNEAPINYNVATDKGVHGLEKLEIQRNNDNDTIIVRECGKCSQQAQCP
ETKLETRKGIYSIYFMGRSVFTGCPQVRTIITRACCAFPFGPOCAQCGRQ
VCSNGFLCDGVNGTGCQCGFNGTACTETGKYGICHDQACSVHGRCSQGLG
DSCDCDVGRGVKDMELTDCNGTCTSANCLLDPGKASCKAAGFRNGTVCT
AINAGETSNGSGSKADCKRTTPGNRVVCVCRAGYTGDDGIVCLEINPCLEHGGCDRNA
ECTQTPNQAVCNCLPKYTGDKVCLNVLNNGGSGPFACNYCTQDQRTCTCKP
DYTGDIYVGRGSIYELPKNESTQYFQLQEHAVRELAPGFTVPFAPLSGSFNEP
RIKWDQOGLMSQVHVGVCQQLLDNLKVTTSATLQGEPSVISVSGYTFVFNNE
AKVLSDDIISTNGVTHVDLKLLSPKLLIIPKDALGRVLONTJVAANHYTKFKLI
ODSGLSVISITNGVTHVDLKLLSPKLLIIPKDALGRVLONTJVAANHYTKFKLI
SKALASDLPRASAKWTLGSELSVRCGTGSDIGELFNEQMCRFIHRGLFDVGVAYG
IDCLMNPITLGGRCDDTFTDIPGEGSCIFTPKCPKSKPKGKKKCIYNPLPFRN
VEGQNLCTVVIQTPRCCHGYFMPDQACPGGPDPTCNRMGMCRDLYTPMGCLCHTG
FNGTACELCHWGRFGPDQPRSCSEHGCGDEGITGSGELCETGTAASCDTPTAVFA
VCTPACSVHATCTENNTCYNLVEGDIITCTVDFCKONNGCAKAKVSKQKQTVS
CSCKGYKGDYSCIEIDPCADGVNGGCHHATCRMTGPGKHCKECSHYVGDVDC
PEOLPLDCLDNGOCHPDASCADLYFDQTTVGVPHLSPGLQYKLTDFDKAEACAKE
AATATYNQLSYAOKAKIHLCSAGLESRVAYPTTASQKCGANVGVIVDYSGRANK
SEMDVFCYRMKDVNCTKAGYVGVGSGNLLQVLSFSLTNFLVFLAFSKSA
RGQAFKLHLDLSIRGTLFVPSNGLPKNKSLGRDIEHLITNVNVSFNDLVNGLFL
RTMGLSOLLITFSQDLHQETFPVDRSILQWDIIAANGKILHIISEPLRAPPTATAA
HSLGTLGTFCAVVLVTGAILAAYSYFRLKORTGFORFDOKRTLMSPASSPRIS
QTLQMBRQRRHPQSPVPTPSQTLNRIWRTAILMWHGCPDMSQATTVTVPR"
1..4293
/mat_peptide
/product="hyaluronan receptor for endocytosis HARE"
/note="175 kDa; derived by proteolytic cleavage"
```

ORIGIN

```
Query Match 59.4%; Score 1936.8; DB 10; Length 4706;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 652; Indels 14; Gaps 6;
QY 11 GGAGCTTCCCAAGAACCCGAAAACTTCCACGATTATTTCTTCCAGTTGCGAGGACATTCGT 70
Db 1395 GGAGCTTCCCAAGAACCCCTTCGACGTCACGACTACTTCTTCCAGTTGCGAGGACATGCTGT 1454
QY 71 GAAAGATCTGTGCGGCCCGAGCCCTTCACTGTGTTTTGACCTTTATCTGAGCCCTTTGA 130
Db 1455 CCGAGAGCTTCTGTGACCTGCGCCCTTCACTGTGTTGCGCCCTTGTCTAGCTCTCTCAA 1514
QY 131 TGAGGAAGCTCGGCTTAAAGACTGGGCAATACGTTTATGCCCCAGGTTCTTCGGTA 190
Db 1515 TCATGAGCCCGGATTAAGACTGGGATCAGCAGGCGCTCATGTCCCAGGTTCTTCGGTA 1574
QY 191 CCATGTGTGCTGCCGCCACAGCTGCTTCTGAAAAACCTGAAATGTCTCAATGCTAC 250
Db 1575 TCAGTGTGTGGCTGCCAGAGCTGCTGTGGCAACCTAAAGTGACCAAGTGCAC 1634
QY 251 TTCCCTCAAGGAGAGCAATAGTATCTCGTCTCTCAGAGCAGGTGTATATAATAA 310
Db 1635 GACCTTCCAGGAGAGCCAGTTTCCATCTCTCTCAGGACACTGTGTTTCATAAACA 1694
QY 311 TAAGGCTAAGATCATATCCAGTGATATCATCATTAATGGGATTTGTCATATCATAGA 370
Db 1695 TGAGGCAAGGTCTGTCCAGTGACATCATCAGCACCAATGGGCGTCATCCACGTTATAGA 1754
QY 371 CAAATTGCTATCTCCCAAAAATTTGCTTATCACTTCCCAAGAGCAAACTCTCTGAAGAAATCT 430
Db 1755 CAAGTGTCTGTCTCCCAAAAATTTGCTTATCACTTCCCAAGAGTCCCTTGGCGAGGTTCT 1814
QY 431 GCAAAATCTTACGACTTTGGCAACAAACATGGCTATCAAAATTTAGCACTTAATATACA 490
Db 1815 GCAAAATCTTACTACAGTGGCAGCAAAACACCGGATATACCAAAATTCAGCAAGTTGATACA 1874
QY 491 GGACTCAGGTTTGTCTGAGTGTCTATCAGGATCCCATCCACACCCAGTCACTCTCTCTG 550
Db 1875 GGACTCAGGTTTGTCTGTCAGTCACTGACTCCATCCACACCCAGTCACTGTCTTCTG 1934
```


Qy	1502	ATGCTCCAGAGGCGCAGAGGTCTCTCTGAGCTGCCAGAGCGATACAAAGGGGACGG	1561
Db	1949	ATGCTCCAGAGGCGCAGAGGTCTCTCTGAGCTGCCAGAGCGATACAAAGGGGACGG	2008
Qy	1562	GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCCTTAACGGAGGGGTGCACGAGCA	1621
Db	2009	GCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCCTTAACGGAGGGGTGCACGAGCA	2068
Qy	1622	CGCCACCTGTAAAGATGACAGGCGCGGGCAAGCAAGTGTGAGTGTAAAGTCACTATGT	1681
Db	2069	CGCCACCTGTAAAGATGACAGGCGCGGGCAAGCAAGTGTGAGTGTAAAGTCACTATGT	2128
Qy	1682	CGGAGATGGGCTGAAGTGTGAGCGGAGCAGCTGCCCATTTGACCGCTGTACAGACAA	1741
Db	2129	CGGAGATGGGCTGAAGTGTGAGCGGAGCAGCTGCCCATTTGACCGCTGTACAGACAA	2188
Qy	1742	TGGGCAAGTGCATGCAGACGCCAAATGTGTGCACCTCCACTTCCAGGATACCACTGTTGG	1801
Db	2189	TGGGCAAGTGCATGCAGACGCCAAATGTGTGCACCTCCACTTCCAGGATACCACTGTTGG	2248
Qy	1802	GGTGTTCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA	1861
Db	2249	GGTGTTCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA	2308
Qy	1862	GGCCTGTGCCAAGAGCTGGACCATGGACCTACACCACTCTCCTATGCCAGAA	1921
Db	2309	GGCCTGTGCCAAGAGCTGGACCATGGACCTACACCACTCTCCTATGCCAGAA	2368
Qy	1922	GGCCAAAGTA	1930
Db	2369	GACCTGGTA	2377

Search completed: June 21, 2004, 14:12:32
Job time : 12320 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 10:23:51 ; Search time 216 Seconds
(without alignments)
8375.648 Million cell updates/sec

Title: US-10-028-248A-210

Perfect score: 3260

Sequence: 1 ggcacgagcaggagctcccc.....aaaaaaaaaaaaaaaaaaaa 3260

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187.4	5.7	1482	4	US-09-016-434-230
2	107.2	3.3	409	4	US-09-833-381-1252
3	81.4	2.5	1414	1	US-08-024-868-1
4	81.4	2.5	1414	2	US-08-242-097-1
5	81.4	2.5	1414	3	US-08-206-695-1
6	81.4	2.5	1414	4	US-09-799-118-1
7	72	2.2	1734	4	US-09-484-970B-63
8	58.6	1.8	1985	4	US-09-907-794A-212
9	58.6	1.8	1985	4	US-09-905-125A-212
10	58.6	1.8	1985	4	US-09-902-775A-212
11	54.2	1.7	7218	1	US-08-232-463-14
12	53.8	1.7	1400	2	US-08-001-078A-2
13	53.8	1.7	1400	2	US-08-463-218-2
14	53.8	1.7	1400	5	PCT-US94-00253-2
15	52.8	1.6	7218	1	US-08-232-463-14
16	47.6	1.5	3259	5	PCT-US95-03747-1
17	44.4	1.4	1587	4	US-09-010-147B-19
18	44.4	1.4	1720	4	US-09-148-545-53
19	43.2	1.3	5191	1	US-08-340-428B-1
20	43.2	1.3	5191	5	PCT-US93-07306-1
21	42.2	1.3	2148	4	US-09-643-597-154
22	42.2	1.3	2148	4	US-09-480-884A-154
23	42.2	1.3	2148	4	US-09-542-615A-154
24	42.2	1.3	2148	4	US-09-606-421B-154
25	42.2	1.3	2148	4	US-09-221-107-154
26	41.8	1.3	619	4	US-09-489-847-58
27	41	1.3	94	4	US-09-621-976-12471

28	41	1.3	8224	6	5180808-1	Patent No. 5180808
29	40.8	1.3	1520	1	US-08-225-477B-1	Sequence 1, Appli
30	40.8	1.3	1520	5	PCT-US95-04353-1	Sequence 1, Appli
c 31	39.6	1.2	832	4	US-09-621-976-2813	Sequence 2813, Ap
32	39.4	1.2	855	6	5185441-40	Patent No. 5185441
33	39.4	1.2	855	6	5223394-3	Patent No. 5223394
34	39.4	1.2	863	1	US-07-940-861-11	Sequence 11, Appl
35	39.4	1.2	863	1	US-08-459-512-11	Sequence 11, Appl
36	39.4	1.2	863	1	US-08-459-657-11	Sequence 11, Appl
37	39.4	1.2	863	2	US-08-460-132-11	Sequence 11, Appl
38	39.4	1.2	863	5	PCT-US92-02050-11	Sequence 11, Appl
39	39.4	1.2	863	6	5185441-35	Patent No. 5185441
40	39.4	1.2	863	6	5223394-5	Patent No. 5223394
41	39.4	1.2	1078	6	5223394-10	Patent No. 5223394
42	39.4	1.2	1519	1	US-08-225-477B-2	Sequence 2, Appli
43	39.4	1.2	1519	5	PCT-US95-04353-2	Sequence 2, Appli
44	39.4	1.2	2007	4	US-09-643-597-153	Sequence 153, App
45	39.4	1.2	2007	4	US-09-480-884A-153	Sequence 153, App

ALIGNMENTS

RESULT 1
US-09-016-434-230
; Sequence 230, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Sealhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 230:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1482 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SPLN004
; CLONE: 1561663
US-09-016-434-230

Query Match 5.7%; Score 187.4; DB 4; Length 1482;
Best Local Similarity 53.2%; Pred. No. 4.3e-45;

	Matches	493;	Conservative	0;	Mismatches	421;	Indels	13;	Gaps	4;			
QY	1659	TGTGAGTGTAA	AGTCACTATGT	CGAGATGG	GCTGAACTGT	---	GAGCCGAG	CAGCTG	1715				
Db	7	TGTGAGTGC	CAACCGCA	GGCTACGT	TAGCGATGG	ACTGCAGTGT	CTGGAGGAGT	CGGAACCA	66				
QY	1716	CCCATTTGAC	CGCTGCTTAC	GAGCAAT	TGGCAGTGC	CAATGCAGAC	CGCAAAATGT	CTGCAC	1775				
Db	67	CTGTGTG	ACCGCTGT	TGGGCC	ACGCA	CCGCCCTGC	CACTCAGAT	CCCATGTG	CACTGAC	126			
QY	1776	CTCCACTT	CCAGGATAC	CACTGT	TGGGTGT	TCCATCTAG	CTCCCACT	TGGGCCAGTAT	1835				
Db	127	CTGCACTT	CCAGAGAA	CGGGCT	TGGGTTTT	CCACTTC	AGGCAC	CACAGCGCCCTTAT	186				
QY	1836	AAGCTCA	CCCTTTGAC	AAAGCC	CAGAGG	CGCTGTGCC	AAAGCT	CGCAATG	GCAACC	1895			
Db	187	GGTCTG	NACTTTTC	GAGGCT	TAGGGCG	CATGCGA	AGAC	-	AGGAGCGCTCT	TGCTTCA	245		
QY	1896	TACAA	CCAGCTT	CTTATG	CCCAAG	AGGC	CAAGTAC	CACTGTG	CTCAGCAG	CTGCTG	1955		
Db	246	TTCCCTC	CAGCTCT	CTGCTG	CCCA	CACTG	GGGCTTCC	ACTGTG	CTCTCA	TGCGGCTGG	305		
QY	1956	GAGAC	GGGGGT	TGCTACT	ACCCAC	ACGCTT	CGCTCC	CAGAACT	GTGGCT	CTGCTGTG	2015		
Db	306	GCCAA	TGGCTC	CACTGCC	CACTGT	GGTTTT	CCCTGT	GCGG	ACTGTG	GCATGTCG	365		
QY	2016	GTGGG	ATAGTGA	CTATG	CACTAG	ACCA	CAAGAG	TGAAAT	TGTTGG	ATGCTTCTGC	2075		
Db	366	GTGGC	ATAGT	CAGCTG	GGTGG	CCGCA	AACTCT	CAGAA	CGTGGG	ATGCTACTGC	425		
QY	2076	TATCG	GAATGA	AGATGA	ACTGCA	CTGCA	AGGTGG	GCTATG	TGGGAGATG	---CTTC	2132		
Db	426	TTCCGT	GTGCA	AGATGT	GGCTG	CCGATG	CGAAAT	TGGCTTC	TGTTGG	TGTCAGCGGATC	485		
QY	2133	TCATG	CAGTGG	GAACCTG	CTGCA	GGTCTCAT	GTCTTCC	CTCACT	CAAACTT	CTCCTG	2192		
Db	486	ACGTG	CAATGG	GAAGCT	GTGGAT	TGCTGG	CTGCCA	CTGCCA	CTTCTC	CACTTCTAT	545		
QY	2193	ACGGA	AGTGTG	GGCTATT	CCAAC	AGCTAG	CTGAG	CGCGTGC	ATTTCT	PAGAACCTG	2252		
Db	546	GGGAT	GCTATT	TGGCTAT	TGCCAAT	GTGCC	ACCCAG	CGGGGTCT	CGACTT	CTCTGG	ACTTCTG	605	
QY	2253	ACTGAC	CTGTCC	ATCCG	GGCACCT	TTTTGT	GCACAG	ACAGTGG	GGCTGG	GGGAGAA	2312		
Db	606	GATGAT	GAGCTC	ACGTATA	AGACACT	CTTCTG	CTCCTGT	CAATGA	AGGCTTT	TGGACA	665		
QY	2313	GAGAC	CTTGTCT	GGGGG	GCACAT	CGAGCAC	CACTCG	CCCAATGT	TCAGCAT	GTTTTCTAC	2372		
Db	666	ATGAC	GTGAT	GTGCCC	CAGACTT	TGGAGT	GCATG	CTCC	ACGCC	ACCCTCTTA	GTAGTGC	725	
QY	2373	AATGAC	CTTGT	CAATG	GCAAC	CCCTG	CAACAG	AGGTGG	GAACAG	CACTGCTCAT	CACT	2432	
Db	726	AAGCC	ACG	CCAGG	GAAGT	TGCTTCC	GGCCCACT	CTCAG	GCCTCAT	CACTCAGT	GTGAC	785	
QY	2433	GCCAG	CCAGG	ACC-----	CACTC	CAACCG	AGGAG	CCAGT	TTGTTG	TATGGA	AGAGCC	2486	
Db	786	GCAGG	CCCTG	CAACAG	TTCT	CTGGG	CCCTGT	TGCCCC	CAGG	ACAGT	TTGTAG	CCGT	845
QY	2487	ATTCTG	CAGTGG	GCATCT	TTTGG	CTCCAA	TGGGAT	CAATCAT	GTTCAT	TTCCAG	GCCTTTA	2546	
Db	846	ATCAT	TGTG	GGGACAT	CATGG	CCCTCA	TGGCAT	CTCC	ATGCT	CTGCG	CACTG	905	
QY	2547	AAAG	ACCCCT	CTG	CCCCG	TG	GA	CCCTG					
Db	906	CTGG	CA	CCCC	CACAG	CCCC	AGG	CA	GTG				

RESULT 2

US-09-833-381-1252
; Sequence 1252, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.

```

; TITLE OF INVENTION: NO. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1252
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(409)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1252

Query Match          3.3%; Score 107.2; DB 4; Length 409;
Best Local Similarity 56.3%; Pred. No. 1.2e-21;
Matches 218; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

QY      1452 GTTGTGGATTTCTGCAACAGCAGCACGGGGGCTGTGCAAAGTGCGCCAGATGCTCCAG 1511
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       23   GTGGCGGANCTGTGCCAAGAAGGGGCATGTGGCTTCAGATTAGCATGCCAACTGCAGCAG 82

QY      1512 AAAGGCAACCAAGGTCTCTGCGACTGCCAAGAGGGGATACAAGGGGACGGGCACAGCTGC 1571
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       83   GTGGGACAGTGTNGTTGCCCTCCCTGCCTGACTATGAGGOTGACGGTGGAGCTGC 142

QY      1572 ACAGATAGACCCCTGTGCGAGCGCCTTAACGAGGGGTGTACGAGCAGCGCACCTGT 1631
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       143 CGAGCTCGAGACCCCTGCTTGGATGGCCACCCTGGAGGTTGCGACGAGCAGCTGATCG 202

QY      1632 AAGATCACAGGCGCGGCAAGCAAGTGTGAGTGTAAAAGTCACTATGTGCGAGATGGG 1691
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       203 CTCATATCCGGTCCGAACACACCGCGCTGTGAATGCCACGTAGGCTACGTGGGTGATGG 262

QY      1692 CTGAATCTGAGCCGGAGC---AGTGCCCATTTGACCGCTGTTCACGACATGGGCAG 1748
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       263 CTCAGTGTCTANAGGAGCTTGAACCCCTGTGGACAGATGCTTGGAGGATCATCACT 322

QY      1749 TGCATGACAGCCCAAAATGTGTCGACCTCCACTTCCAGGATACCACTGTGGGGTGTTC 1808
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       323 TGCCACACTGATGCTTTGTGCATGACCTACATTTCCAGAAAAACAGGCTGGTGTCTTC 382

1809 CATCTAGCTCCCACCTGGGCCAGTAT 1835
383 CACATCCAGGCCACCACTGGCCCTTAT 409

```

RESULT 3

```

US-08-024-868-1
; Sequence 1, Application US/08024868
; Patent No. 5386013
; GENERAL INFORMATION:
; APPLICANT: Lee, Tae Ho
; APPLICANT: Wisniewski, Hans Georg
; APPLICANT: Vilcek, Jan
; TITLE OF INVENTION: Cytokine-Induced
; TITLE OF INVENTION: Coding Therefor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: US
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA: US/08/024,868
APPLICATION NUMBER: US/08/024,868
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/642,312
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Livnat, Shmuel
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: VILCEK=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-628-5197
TELEFAX: 212-737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1414 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Fibroblast
CELL LINE: FS-4
FEATURE:
NAME/KEY: CDS
LOCATION: 69..899
OTHER INFORMATION:
US-08-024-868-1

Query Match 2.5%; Score 81.4; DB 1; Length 1414;
Best Local Similarity 55.5%; Pred. No. 1.3e-13;
Matches 157; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
QY 1795 CTGTGGGGTGTTCATCTACGCTCCCGACCTGGGCGCAGTATAGCTGACCTTTGACAAAG 1854
Db 169 CAGCGGTGTGTACACAGAGACGCGTCTGGCAATACAGCTACCTACCGAGAG 228
QY 1855 CCAGAGAGCGCTGTGCCAAGAGCTGCGACCATGCGCAACCTACAAACAGCTCTCCTATG 1914
Db 229 CTAAGCGGTGTGTGAATTTGAAGCGCGCCATCTCGCAACTTACAGCAGCTAGAGCAG 288
QY 1915 CCCAGAGAGCGCAAGTACACCTGTGCTCAGCAGCTGCTGGAGACCGGGCGGTTCCT 1974
Db 289 CCAGAAAATTTGGATTTTCATGCTGTGCTGATGGATGGCTAAAGGCGAGAGTTGGAT 348
QY 1975 ACCCCACAGCCTTCGCCCTCCAGAACTGTGCTCTGTGTTGGGATAGTGACTATG 2034
Db 349 ACCCATTTGTGAAGCCAGGCGCCCACTGTGATTTGAAAACCTGGCATTTATTGATTATG 408
QY 2035 GACCTAGACCCCAACAGAGTGAATGTGGGATGCTTCTGCTA 2077
Db 409 GAATCCGCTCTCAATAGGAGTGAAGATGGGATGCTATTGCTA 451

RESULT 4
US-08-242-097-1
Sequence 1, Application US/08242097
Patent No. 5846763
GENERAL INFORMATION:
APPLICANT: Lee, Tae Ho
APPLICANT: Wisniewski, Hans Georg
APPLICANT: Vilcek, Jan
TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
TITLE OF INVENTION: Therefor and Uses Thereof
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.

ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/242,097
FILING DATE: 13-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: LEE26/VILCEK=1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-628-5197
TELEFAX: 212-737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1414 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Fibroblast
CELL LINE: FS-4
FEATURE:
NAME/KEY: CDS
LOCATION: 69..899
OTHER INFORMATION:
US-08-242-097-1

Query Match 2.5%; Score 81.4; DB 2; Length 1414;
Best Local Similarity 55.5%; Pred. No. 1.3e-13;
Matches 157; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
QY 1795 CTGTGGGGTGTTCATCTACGCTCCCGACCTGGGCGCAGTATAGCTGACCTTTGACAAAG 1854
Db 169 CAGCGGTGTGTACACAGAGACGCGTCTGGCAATACAGCTACCTACCGAGAG 228
QY 1855 CCAGAGAGCGCTGTGCCAAGAGCTGCGACCATGCGCAACCTACAAACAGCTCTCCTATG 1914
Db 229 CTAAGCGGTGTGTGAATTTGAAGCGCGCCATCTCGCAACTTACAGCAGCTAGAGCAG 288
QY 1915 CCCAGAGAGCGCAAGTACACCTGTGCTCAGCAGCTGCTGGAGACCGGGCGGTTCCT 1974
Db 289 CCAGAAAATTTGGATTTTCATGCTGTGCTGATGGATGGCTAAAGGCGAGAGTTGGAT 348
QY 1975 ACCCCACAGCCTTCGCCCTCCAGAACTGTGCTCTGTGTTGGGATAGTGACTATG 2034
Db 349 ACCCATTTGTGAAGCCAGGCGCCCACTGTGATTTGAAAACCTGGCATTTATTGATTATG 408
QY 2035 GACCTAGACCCCAACAGAGTGAATGTGGGATGCTTCTGCTA 2077
Db 409 GAATCCGCTCTCAATAGGAGTGAAGATGGGATGCTATTGCTA 451

RESULT 5
US-09-206-695-1
Sequence 1, Application US/09206695
Patent No. 6210905
GENERAL INFORMATION:
APPLICANT: Lee, Tae Ho
APPLICANT: Wisniewski, Hans Georg
APPLICANT: Vilcek, Jan
TITLE OF INVENTION: Cytokine-Induced Protein, TSG-6, DNA Coding
TITLE OF INVENTION: Therefor and Uses Thereof
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W., Suite 300

RESULT 7
US-09-484-970B-63
; Sequence 63, Application US/09484970B

```

RESULT 6
US-09-799-118-1
: Sequence 1, Application US/09799118
: Patent No. 6518401
: GENERAL INFORMATION:
: APPLICANT: Lee, Tae Ho
: Wisniewski, Hans
: Vilcek, Jan
: TITLE OF INVENTION: Cytokine
:

```



```
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 63
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6426186 201752.1
; LOCATION: 1444, 1455, 1457, 1461, 1463, 1465, 1467, 1497, 1500, 1506, 1521, 1534-
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-63

Query Match          2.2%; Score 72; DB 4; Length 1734;
Best Local Similarity 55.6%; Pred. No. 9.2e-11;
Matches 158; Conservative 0; Mismatches 125; Indels 1; Gaps 1;

QY 1795 CTGTTGGGGTTCATCTACGCTCCCACTGGCCAGTATAGCTGACCTTTGACAAG 1854
DB 169 CAGCGGTGTGTACACAGAGACGCGTCTGGCAATACAACTACCTACGCGAAG 228
QY 1855 CCAGAGGCGCTGTCCCAAGAGCTGCACCATGGCAACCTACAAAGCTCTCTATG 1914
DB 229 CTAAGCGGTGTGAAATTGAAGCGCGCCACTCTCCCACTTACAAAGCAGCTAGAGCGAG 288
QY 1915 CCCAAGGCCCAAGTACCACCTGTCTCAGCAGGCTGGCTGGAGACCGGGGTTGCT 1974
DB 289 CCAGAAATTTGGATTTTCATCTCTGTCTGTGGATGATGCTAAGGCGAGATTGGAT 348
QY 1975 ACCCAACAG-CCTTCGCTCCCAAGTGTGGCTCTGGTGTGGTGGATAGTGGACTAT 2033
DB 349 ACCCATTTGTGAAGCGAGCGGCCCACTGTGGATTGGAAAACTGGCATTTATTGATT 408
QY 2034 GGACCTACACCAACAAGAGTGAATGTGGGATGCTTCTGCTA 2077
DB 409 GGAATCCGCTCTCATAGGATGAAGATGGGATGCTATTGCTA 452

RESULT 8
US-09-907-794A-212
; Sequence 212, Application US/09907794A
; Patent No. 6635468
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
```

```
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-794A-212

Query Match          1.8%; Score 58.6; DB 4; Length 1985;
Best Local Similarity 55.0%; Pred. No. 9.6e-07;
Matches 115; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1797 GTTGGGTGTTCATCTACGCTCCCACTGGCCAGTATAGCTGACCTTTGACAAGCC 1856
DB 593 GGTGTGTCTTCTCTTACCAGTCCCAACCGGCGCTACCACTTCACTCCACGAGGCG 652
QY 1857 AGAGAGCGCTGTGCCAAGAGCTGCCAGCATGCAACCTACACAGCTCTCTCTATGCC 1916
DB 653 CAGCAGGTCTGTGCAGAGCAGGCTGCGGTGGTCTCTCTTTGAGCAGCTCTTCCGGGCC 712
QY 1917 CAGAAGGCCAAGTACCACTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGTTGCCCTAC 1976
DB 713 TGGAGAGAGGCGCTGGACTGTGTGCAACCGGCGCTGGCTGCGATGTACGGTGCAGTAC 772
QY 1977 CCCACAGCTTTCGCTCTCCCAAGAACTGTGG 2005
DB 773 CCCATCATGTTGCCCGCGCAGCCCTGCGG 801

RESULT 9
```

US-09-905-125A-212

; Sequence 212, Application US/099505125A
 ; Patent No. 6664376
 : GENERAL INFORMATION:

APPLICANT: Genentech, Inc.	PCT/US99/21547
APPLICANT: Ashkenazi, David	PCT/US99/23089
APPLICANT: Bostein, David	PCT/US99/23089
APPLICANT: Desnoyers, Luc	PCT/US99/23089
APPLICANT: Eaton, Dan L.	PCT/US99/23089
APPLICANT: Ferrara, Napoleone	PCT/US99/23089
APPLICANT: Filvaroff, Ellen	PCT/US99/23089
APPLICANT: Fong, Sherman	PCT/US99/23089
APPLICANT: Gao, Wei-Qiang	PCT/US99/23089
APPLICANT: Gerber, Hanspeter	PCT/US99/23089
APPLICANT: Gerritsen, Mary E.	PCT/US99/23089
APPLICANT: Goddard, A.	PCT/US99/23089
APPLICANT: Godowski, Paul J.	PCT/US99/23089
APPLICANT: Grimaldi, Christopher J.	PCT/US99/23089
APPLICANT: Gurney, Austin L.	PCT/US99/23089
APPLICANT: Hillan, Kenneth, J.	PCT/US99/23089
APPLICANT: Klijavin, Ivar J.	PCT/US99/23089
APPLICANT: Mather, Jennie P.	PCT/US99/23089
APPLICANT: Pan, James	PCT/US99/23089
APPLICANT: Paoni, Nicholas F.	PCT/US99/23089
APPLICANT: Roy, Margaret Ann	PCT/US99/23089
APPLICANT: Stewart, Timothy A.	PCT/US99/23089
APPLICANT: Tumas, Daniel	PCT/US99/23089
APPLICANT: Williams, P. Mickey	PCT/US99/23089
APPLICANT: Wood, William, I.	PCT/US99/23089
TITLE OF INVENTION: Secreted and Transmembrane	
TITLE OF INVENTION: Acids Encoding the S	
FILE REFERENCE: 10466-14	
CURRENT APPLICATION NUMBER: US/09/905,129	
CURRENT FILING DATE: 2001-07-12	
PRIOR APPLICATION NUMBER: PCT/US00/04414	
PRIOR FILING DATE: 2000-02-22	
PRIOR APPLICATION NUMBER: US 60/143,048	
PRIOR FILING DATE: 1999-07-07	
PRIOR APPLICATION NUMBER: US 60/145,698	
PRIOR FILING DATE: 1999-07-26	
PRIOR APPLICATION NUMBER: US 60/146,222	
PRIOR FILING DATE: 1999-07-28	
PRIOR APPLICATION NUMBER: PCT/US99/20594	
PRIOR FILING DATE: 1999-09-08	
PRIOR APPLICATION NUMBER: PCT/US99/20944	
PRIOR FILING DATE: 1999-09-13	
PRIOR APPLICATION NUMBER: PCT/US99/21090	
PRIOR FILING DATE: 1999-09-15	
PRIOR APPLICATION NUMBER: PCT/US99/21547	
PRIOR FILING DATE: 1999-09-15	
PRIOR APPLICATION NUMBER: PCT/US99/23089	
PRIOR FILING DATE: 1999-10-05	
PRIOR APPLICATION NUMBER: PCT/US99/28214	
PRIOR FILING DATE: 1999-11-29	
PRIOR APPLICATION NUMBER: PCT/US99/28313	
PRIOR FILING DATE: 1999-11-30	
PRIOR APPLICATION NUMBER: PCT/US99/28564	
PRIOR FILING DATE: 1999-12-02	
PRIOR APPLICATION NUMBER: PCT/US99/28565	
PRIOR FILING DATE: 1999-12-02	
PRIOR APPLICATION NUMBER: PCT/US99/30095	
PRIOR FILING DATE: 1999-12-16	
PRIOR APPLICATION NUMBER: PCT/US99/30911	
PRIOR FILING DATE: 1999-12-20	
PRIOR APPLICATION NUMBER: PCT/US99/30999	
PRIOR FILING DATE: 1999-12-20	
PRIOR APPLICATION NUMBER: PCT/US00/00219	
PRIOR FILING DATE: 2000-01-05	
NUMBER OF SEQ ID NOS: 423	

US-09-905-125A-212

	Query Match	1.8%;	Score 58.6;	DB 4;	Length 1985;
	Best Local Similarity	55.0%;	Pred. No. 9.6e-07;		
	Matches 115;	Conservative	0;	Mismatches 94;	Indels 0; Gaps 0;
Qy	1797	GTGCGGCGTTCACCTCTACGCTCCCCACATGGGCGAGTATAGCTGACCTTTGACAAAGCC	1856		
Db	593	GGTGTGGTCTTTCCTTACCAGTCCCCCAACGGGCGCTACCAAGTTCAACTTCCACGAGGGC	652		
Qy	1857	AGAGAGCCTGTGCCAACGAGAGCTGGCCACATGGCAACCTACACAGCTCTCTCTATGCC	1916		
Db	653	CAGCAGCTCTGTGCAGACGAGGCTGGCGGTGGTGGCCCTCCTTTGACAGCTCTTCCGGGCC	712		
Qy	1917	CAGAAGGCCAAGTACCAACTGTCTCAGCAGGCTTGCTCGAGACCGGGCGGTTGCTCTAC	1976		
Db	713	TGGAGAGAGGCCCTGACTGCTGTCACAGCGGGCTGGCTGCAGGATCTACGGTGCAGTAC	772		
Qy	1977	CCACAGCCTTCGCTCCCGCAAACTGTGG	2005		
Db	773	CCCATCATGTGGCCCGGCGAGCCCTCGG	801		

RESULT 10

US-09-902-775A-212
: Sequence 212, Application US/09902775A

; Patent No. 6686451

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Ashkenazi, Avi

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Filvarzoff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, A.

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth, J.

APPLICANT: Klijavin, Ivar J.

APPLICANT: Mather, Jennie P.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William, I.

TITLE OF INVENTION: Secured and Tra

TITLE OF INVENTION: Acids Encoding

FILE REFERENCE: 10466-14

CURRENT APPLICATION NUMBER: US/09/30

CURRENT FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: PCT/US00/07

PRIOR FILING DATE: 2000-02-22

PRIOR APPLICATION NUMBER: US 60/143,

PRIOR FILING DATE: 1999-07-07

PRIOR APPLICATION NUMBER: US 60/145,

PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: US 60/146,

PRIOR FILING DATE: 1999-07-28

PRIOR APPLICATION NUMBER: PCT/US99/2

PRIOR FILING DATE: 1999-09-08

PRIOR APPLICATION NUMBER: PCT/US99/2

PRIOR FILING DATE: 1999-09-13

PRIOR APPLICATION NUMBER: PCT/US99/2

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/2

```
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-775A-212

Query Match      1.8%; Score 58.6; DB 4; Length 1985;
Best Local Similarity 55.0%; Pred. No. 9.6e-07;
Matches 115; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1797 GTTGGGTTCCATCTACGCTCCCACTGGCCAGTATAGCTACCTTTGACAAAGCC 1856
Db 593 GGTGTGTTCTTCTTACAGTCCCAAGCGGCGCTACCAAGTCAACTCCACGAGGC 652
QY 1857 AGAGAGGCTGTGCCAAGAGCTGCGACCATGCGCAACCTACCAACCACTCTCTATGCC 1916
Db 653 CAGCAGGTCTGTGAGAGCAGCTGCGGTGTGGCTCTTTGACGAGCTTTCCGGGCC 712
QY 1917 CAGAAGGCAAGTACCACTGTGTGACAGGCTGGTGGAGACGGCGCGGTTGCCCTAC 1976
Db 713 TGGAGAGGAGGCGCTGGACTGTGTGAACCGCGGCTGGCTGCAGGATGCTACGGTGCAGTAC 772
QY 1977 CCACAGCCTTGGCTCCCACTCCCAACTGTGG 2005
Db 773 CCATCATGTTGCCCGCAGCCCTGCGG 801

RESULT 11
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match      1.7%; Score 54.2; DB 1; Length 7218;
Best Local Similarity 3.6%; Pred. No. 4.6e-05;
Matches 14; Conservative 221; Mismatches 154; Indels 0; Gaps 0;

QY 1392 AAGGAGAACACACGCTGTGAGTGAACCTGATTATGAAGTGAAGGAGGATCACATGCACA 1451
Db 1432 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1373
QY 1452 GTTGTGATTTCGCAACACAGACACAGGGGGCTGTGCAAGGTGGCCAGATGTCCTCCAG 1511
Db 1372 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1313
QY 1512 AAGGCGACGAGGTCTCTGACAGTCCAGAGGATACAAAGGGACGGGACAGCTGC 1571
Db 1312 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1253
QY 1572 ACAGAGATAGACCCCTGTGACAGACGGCTTAACGAGGGGTGTACAGGACGCCACCTGT 1631
Db 1252 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1193
QY 1632 AAGATGACAGCCCGGCAAGACAGTGTAGTGTAAAGTCACTATGTCGAGATGGG 1691
Db 1192 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1133
QY 1692 CTGAACTGTAGCGGAGAGCTGCCCATTCGCTTACAGGCAATGGGACGTGC 1751
Db 1132 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1073
QY 1752 CATGCAGACGCCAAATGTGTCGACCTCCA 1780
Db 1072 RRRRRRATCGAAGCTCCCTCGACCTGCA 1044
```

```
RESULT 12
US-08-001-078A-2
; Sequence 2, Application US/08001078A
; Patent No. 5872094
; GENERAL INFORMATION:
; APPLICANT: Goetinck, Paul F.
; APPLICANT: Tondravi, M., Mehrdad
; APPLICANT: Binette, Francois
; TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
; TITLE OF INVENTION: FORMATION
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: Massachusetts
; COUNTRY: USA
```

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/001,078A
FILING DATE: 06-JAN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-008
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1400 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-001-078A-2

Query Match 1.7%; Score 53.8; DB 2; Length 1400;

Best Local Similarity 53.6%; Pred. No. 2e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 112; Conservative 0

QY 1797 GTTGGGGTGTCCATCTACGCTCCCAAGCTGGCCAGTATAGCTGACCTTTGACAAAGCC 1856
DB 706 GGTGTGGTATTCCTTACTTCCACGACTGGGGCGCTACAATCTCAATTTTCACGAGGCG 765
QY 1857 AGAGAGGCGCTGTGCCAAGCAAGCTGCACCTGGCCAGTATAGCTGACCTTTGACAAAGCC 1916
DB 766 CAGCAGGCGCTGTCCAGCAGGATGCTGTGATCGCTCTTCCGACGAGCTGACGAGCGCC 825
QY 1917 CAGAGGCGCTGTGCCAAGCAAGCTGCACGAGCTGGCTGAGAGCGGGGGGTGCTTAC 1976
DB 826 TGGCGGGGGCGCTGGAGCTGGTGGTCAATGCCGCTGGCTCAGTGATGGCTCTGTGCAATCT 885
QY 1977 CCCACAGCCTTCGCTCCCAAGCAAGCTGGCTGG 2005
DB 886 CCCATCAAAAGCCAGAGAGCCCTGTGG 914

RESULT 13

US-08-463-218-2
Sequence 2, Application US/08463218
Patent No. 5986052
GENERAL INFORMATION:
APPLICANT: Goetnick, Paul F.
APPLICANT: Tondravi, Mehrdad
APPLICANT: Binette, Francois
TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, Suite 510
CITY: BOSTON
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,218
FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/001,078
FILING DATE: 06-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-008DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1400 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-463-218-2

Query Match 1.7%; Score 53.8; DB 2; Length 1400;

Best Local Similarity 53.6%; Pred. No. 2e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 112; Conservative 0

QY 1797 GTTGGGGTGTCCATCTACGCTCCCAAGCTGGCCAGTATAGCTGACCTTTGACAAAGCC 1856
DB 706 GGTGTGGTATTCCTTACTTCCAGCAGTGGGGCGCTACAATCTCAATTTTCACGAGGCG 765
QY 1857 AGAGAGGCGCTGTGCCAAGCAAGCTGCACCTGGCCAGTATAGCTGACCTTTGACAAAGCC 1916
DB 766 CAGCAGGCGCTGTCCAGCAGGATGCTGTGATCGCTCTTCCGACGAGCTGACGAGCGCC 825
QY 1917 CAGAGGCGCTGTGCCAAGCAAGCTGCACGAGCTGGCTGAGAGCGGGGGGTGCTTAC 1976
DB 826 TGGCGGGGGCGCTGGAGCTGGTGGTCAATGCCGCTGGCTCAGTGATGGCTCTGTGCAATCT 885
QY 1977 CCCACAGCCTTCGCTCCCAAGCAAGCTGG 2005
DB 886 CCCATCAAAAGCCAGAGAGCCCTGTGG 914

RESULT 14

PCT-US94-00253-2
Sequence 2, Application PC/TUS9400253
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
NUMBER OF SEQUENCES: 2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00253
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,078
FILING DATE: 06-JAN-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1400 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US94-00253-2

Query Match 1.7%; Score 53.8; DB 5; Length 1400;

Best Local Similarity 53.6%; Pred. No. 2e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 112; Conservative 0

QY 1797 GTTGGGGTGTCCATCTACGCTCCCAAGCTGGCCAGTATAGCTGACCTTTGACAAAGCC 1856
DB 706 GGTGTGGTATTCCTTACTTCCAGCAGTGGGGCGCTACAATCTCAATTTTCACGAGGCG 765
QY 1857 AGAGAGGCGCTGTGCCAAGCAAGCTGGCAGCTGAGAGCGGGGGGTGCTTCTCTATGCC 1916

Db 766 CAGCAGGCTGTCTGAGCAGAGTCTGTGATCGCCTCTTCGACCACTGTACGACGCC 825
Qy 1917 CAGAAGGCAAGTACCACTGTCTGCTCAGCAGCTGGCTGGAGACCGGGCGGGTTGCCCTAC 1976
Db 826 TGGCGGGCGGCTGGACTGTGTGAATCGCGCTGGCTCAGTGATGGCTCTGTGCAATCT 885
Qy 1977 CCCACAGCTTCGGCTCCCAAGACTGTGG 2005
Db 886 CCCATCAAAAGCCAGAGAGCCCTGTGG 914

RESULT 15

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ptzgt-Fls
US-08-232-463-14

Query Match 1.6%; Score 52.8; DB 1; Length 7218;
Best Local Similarity 2.7%; Pred. No. 0.00012;
Matches 9; Conservative 198; Mismatches 126; Indels 0; Gaps 0;
Qy 2870 TGGACGGGAGATGCCAGCATCACTCACTGCCACCTGGGCCATCACTGTAATTCAG 2929
Db 1050 TCGAGGAGCTTGGATTT 1109
Qy 2930 CACCAGTTGCGCTTTTAGAGTAAAGTCTTTAAGCACTCAGAACCATACCTCATCTC 2989
Db 1110 YY 1169

Qy 2990 TCTGGCTGATCTGGGGGTGTGTTTCTGTGGGTGAGAGATGTGTGCTGCCCCACCCAGTA 3049
Db 1170 YY 1229
Qy 3050 CAGCTTCCTCTCTGACCCCTTTGGCTCTTCTCCTTTGTAATCTTTCAGCTGGCAGCTGCT 3109
Db 1230 YY 1289
Qy 3110 CCATTCTGCCCTACATGATGGTAACCTGATCTTCTTCCTCCTGTAGATTGTAGACCTC 3169
Db 1290 YY 1349
Qy 3170 CNTCTTGTATCCAGCCCTTAGCCAGTGCCT 3202
Db 1350 YY 1382

Search completed: June 21, 2004, 16:27:07
Job time : 222 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 10:16:51 ; Search time 7854 Seconds

(without alignments)

12395.048 Million cell updates/sec

Title: US-10-028-248A-210

Perfect score: 3260

Sequence: 1 ggcacgacgaggtctccc.....aaaaaaaaaaaaaaaaaaaa 3260

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estro:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Length	ID	Description
1	1852.6	56.8	3085	11 AK034522
2	735.8	22.6	741	14 CA449310
3	591	18.1	605	14 CA433190
4	510	15.6	524	13 BX101101
				EX101101 BX101101

5	486	14.9	562	12	BG220572	BG220572 RST40358
6	472	14.5	481	9	AI081595	AI081595 ox63h05.s
7	441.4	13.5	455	9	AI248408	AI248408 qh65e09.x
8	418.4	12.8	421	9	AI131339	AI131339 qcl8c08.x
9	413.6	12.7	664	10	BB626584	BB626584 BB626584
10	404.2	12.4	488	14	R92610	R92610 yq07g08.s1
11	401.4	12.3	422	9	AA968560	AA968560 og75d11.s
12	381	11.7	448	14	H49176	H49176 yq18g02.r1
13	378.2	11.6	484	9	AA010567	AA010567 z109b11.s
14	370.4	11.4	3025	11	BC049247	BC049247 Mus muscu
15	358.4	11.0	396	14	H91274	H91274 yu99e07.s1
16	355	10.9	819	13	BCX71715	BCX71715 BX71715
17	333	10.2	334	9	AI084842	AI084842 ow89e11.s
18	332.4	10.2	445	14	H90364	H90364 yu99e07.r1
19	329.8	10.1	459	14	T47504	T47504 yb14f01.r1
20	329	10.1	453	14	H49088	H49088 yq18f08.s1
21	328.8	10.1	545	28	AQ278586	AQ278586 CITBI-E1-
22	326	10.0	331	14	R97435	R97435 yq53h07.s1
23	323	9.9	324	9	AA704436	AA704436 z122h04.s
24	322.6	9.9	414	14	H70480	H70480 Yt91e04.s1
25	322.2	9.9	460	14	R97483	R97483 Yq53h06.r1
26	314	9.6	325	9	AA257161	AA257161 zr82b05.r
27	313.2	9.6	1155	12	EM919585	EM919585 AGENCOCURT
28	305.4	9.4	394	9	AA004325	AA004325 zh91f02.s
29	301	9.2	302	9	AA257068	AA257068 zr82b05.s
30	300.8	9.2	2368	11	BC019712	BC019712 Mus muscu
31	286.4	8.8	694	14	CA376384	CA376384 654727 NC
32	283.6	8.7	683	12	BJ523552	BJ523552 BJ523552
33	273.2	8.4	348	14	R91674	R91674 Yq10f09.r1
34	270.8	8.3	328	14	R91675	R91675 Yq10f09.s1
35	269.8	8.3	793	12	B1832486	B1832486 603082278
36	247.8	7.6	385	10	BB842118	BB842118 BB842118
37	240	7.4	601	14	CB513007	CB513007 ssalr9b54
38	216.4	6.6	228	14	T47505	T47505 Yb14f01.s1
39	208.4	6.4	500	9	AA821602	AA821602 vw30c03.r
40	208	6.4	722	13	BCX739778	BCX739778 BX739778
41	206.8	6.3	462	10	BE56181	BE56181 UI-M-BH0-
42	203	6.2	809	9	AU080852	AU080852 AU080852
43	199	6.1	928	12	BI906024	BI906024 603062355
44	194.6	6.0	660	13	EX265613	EX265613 BX265613
45	192	5.9	612	13	EX265614	EX265614 BX265614

ALIGNMENTS

RESULT 1

AK034522

LOCUS

DEFINITION

AK034522

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AK034522 3085 bp mRNA linear HTC 18-SEP-2003
Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:9330210123 product:similar to CD44-LIKE
PRECUSOR FELL [Homo sapiens], full insert sequence.

AK034522 GI:26330004

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

REFERENCE AUTHORS	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	20330913 11076861 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
REFERENCE AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
TITLE	6 (bases 1 to 3085)
JOURNAL AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
TITLE	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/.
COMMENT	Location/Qualifiers 1. 3085 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:9330210123" /db_xref="MGI:2398728" /db_xref="taxon:10090" /clone="9330210123" /sex="male" /tissue type="diencephalon" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev stage="adult" 1..2685 /note="unnamed protein product; putative similar to CD44-LIKE PRECURSOR FELL [Homo sapiens] (SPTR Q9NRY3, evidence: FASTA, 78.1%ID, 99.4%length, match=2673)" /codon_start=1 /protein_id="BAC28741.1"
FEATURES source	1. 3085 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:9330210123" /db_xref="MGI:2398728" /db_xref="taxon:10090" /clone="9330210123" /sex="male" /tissue type="diencephalon" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev stage="adult" 1..2685 /note="unnamed protein product; putative similar to CD44-LIKE PRECURSOR FELL [Homo sapiens] (SPTR Q9NRY3, evidence: FASTA, 78.1%ID, 99.4%length, match=2673)" /codon_start=1 /protein_id="BAC28741.1"
CDS	<1..2685 /note="unnamed protein product; putative similar to CD44-LIKE PRECURSOR FELL [Homo sapiens] (SPTR Q9NRY3, evidence: FASTA, 78.1%ID, 99.4%length, match=2673)" /codon_start=1 /protein_id="BAC28741.1"
translation="GI:26330005"	
AKVLSDDIISNGVIHVIDTLSPONLLITPKGASRVLLNLTVAANHGTFFKSKLI QDSGLLVITDPMTPTVLPWPTDKALQALPOEQDDFLFNDNDKKYALKLPHVIRD TMAUASDLPRASWKTILQGSLSVRCGTSVGLFPLNGQCRILQRELLPDGGVAYG IDCLMPTEGGRCDTFTFNIPEGSCSFSTPRCPLOSKEGKVKKICINPLPRRD VEGQNLCTLVHVRPCSGYFMPDCAQPGDPITPCNNRCMGCDYQKPTQCQCCHTG PNGTACELCLPRFPGPCQPCGSEHQCDEGITGSGCLCEAGTWGFRCDAPTVP VICIPACSMHATCMENNTVCNLYEGDGLTCTVDFCKONNGGCAKAKVKSQKQVVS CSCOKGYKGDHSCSTEDPCANGVGGCHEATCMTCGPKGCKECKSHYVGDGRDCE PEQLDRCLQDNGOCHDPCANCVLHDPDITVGVPHLSPLGQYKLTFDKAKEACE AASATATNQLSYAKAKYKHLCSAGLWESGRVAYPTIYASKKCANILVIVDGTNRKS EMWDVFCYRMKDQVNTCKAGYVGGDFSCNLLQVLMFSLTINFLTIVFVSRSAQ GRFLKHLTDLISITGLFVQNSGLPKNSLGRDIIEHLTNVNVFDDLVNGTVLK TRGLQLLITSSQDLHQEAFVDRALOWDIIASNGVLHIIISPLKAPPTAAAH SGLGTGFCAVLVLTGALAAAYFYRLNORTTGRPRESEDDIDALAFKQQPESIT NPLVETSTPAAPFSCDPTDSGRRELENSDPLGLRS"	
3068..3073 /note="putative"	
3085 /note="putative"	
polyA_signal	
polyA_site	
ORIGIN	
Query Match	56.8%; Score 1852.6; DB 11; Length 3085;
Best Local Similarity	78.3%; Pred. No. 0;
Matches 2333; Conservative	0; Mismatches 614; Indels 32; Gaps 8;
QY	186 CGGTACCATGTCGTCGCCACGAGTCCTTCTGGAAAACCTGAAATGATCTCAAAAT 245
Db	
QY	4 CGGTATCATGTCGTCGCCACGAGTCCTTCTGGAAAACCTGAAATGATCTCAAAAT 63
Db	
QY	246 GCTACTTCCCTCCAAAGGAGAGCAATAGTTCATCTCGTCTCTCAGACGACGGTGTATATA 305
Db	
QY	64 GCCACAACTCCAAAGGAGAGCAATTTCCATCTCTGTCCTCAGGACACTGTGCTCATATC 123
Db	
QY	306 AATAAAGGCTAAGATCATATCCAGTGTATATCATCATGACTACTAATGGGATTTGTTATATC 365
Db	
QY	124 AACCAAGAGGCGAAGGTTCTGTCAGTGCATATATCAGCACCAACGAGGTCATCCACGTC 183
Db	
QY	366 ATAGACAAATTCGTATCTCCAAAATTTGCTTATCATCTCCAAAAGACAACTCTGGAAGA 425
Db	
QY	184 ATAGACAAATTCGTATCTCCAAAATTTGCTTATCATCTCCAAAAGGTCCTCCGGCAGG 243
Db	
QY	426 ATTCTGCAAAATTCGTATCTCCAAAATTTGCTTATCATCTCCAAAAGGTCCTCCGGCAGG 485
Db	
QY	244 GTTCTGCTAAACCTTACTACAGTGGCAGCAACCAATGGATATATCCAAATTCAGCAAGTTG 303
Db	
QY	486 ATACAGGACTCAGGTTTGTGAGTGTCTATCATCCGATCCCATCCACACCCGAGTCACTCTC 545
Db	
QY	304 ATACAGGACTCAGGTTTGTGAAAGTTCATCTGACCCCATGATACCCCGAGTCACTCTC 363
Db	
QY	546 TTCTGGCCACCGACCAAGCCCTCCATGCCCTACCTGCTGAAACAAAGGACTTCTCTGTTTC 605
Db	
QY	364 TTCTGGCCCTACGGACAAAGCCCTCCAGGCTTTCCTCAGGAGCAGCAGACTTCTCTGTTTC 423
Db	
QY	606 AACCAAGACCAACAAAGGACCAAGCTGGAAGGAGTATTTGAAGTTTCAATGTGATACGAGTCC 665
Db	
QY	424 ANTGAAGACCAACAAAGGACCAAGCTGGAAGGATACCTCAAGTTCCACGTCGATCCGAGATAC 483
Db	
QY	666 AAGGTTTGTAGCTGTGATCTTCCACATCCACTGCTGGAAGACCCCTGCAAGGTTTCAGAG 725
Db	
QY	484 ATGGCTTTTAGCTTTCAGACCTCCCGAGGTTCTGCTTCTGGAAGACCCCTGCAAGGTTTCGAG 543
Db	
QY	726 CTGAGTGTGAAATGTGAGGCTGGCAGGACATCGGTGACCTCTTCTCAATGGCCAAACC 785
Db	
QY	544 CTGAGTGTGAGGTTGTGAACTGGCAGTGTATGTCGGAGAGCTCTTCTTAAATGGCAAAATG 603
Db	
QY	786 TGCAGAAATGTGACGCGGGAGCTCTTGTGTTGACCTGGGTGGCCCTACCGCATTCAGTGT 845
Db	
QY	604 TGCAGGATTTACAGCGGAGACTCTTGTGTTGACGGAGGTTGTGCCCTATGGCAATTCAGTGT 663
Db	
QY	846 CTGCTGATTCATCCACCTGGGGGCGCTGTGACACCTTTACTACTTTTCGATCGCTCG 905
Db	

Db 664 CTGCTAATGATCCTACCAAGGTGGCCGATGTGACACTTTTACTACCTTCAATATTCGG 723
QY 906 GGGAGTGTGGAGCTGTGTCAATATCTCCAGCTGCCAAAGGTGGAGTAAACCAAAGGT 965
Db 724 GGGGAATGTGGAGCTGTGTTCCTCACTCCAGATGCCACTACAGACAAACCAAAGGT 783
QY 966 GTGAAGCAGAGTGTCTCTACAACTG---CCCTTCAGAGAGAACTTGGAGGCTGCCGG 1022
Db 784 GTGAGGAAGAGTGTATCTACAAACCGTTACCTTTTCAGAGGGATGTGGAAGCTGTGAG 843
QY 1023 GAGCGGTGAGCTGTGTGATACAGATCCCAAGTGTCTGCAAGGGCTACTTCGGGGAGAC 1082
Db 844 AACCTGTGACACCTGTGTGTCATGTCCGCGGTGTGAGTGGTATCTTATGCGACAG 903
QY 1083 TGTGAGGCTGTGCTGTGAGGACAGATGCCCGTGTATTAACCGGGGTGTCTGCCCTTGT 1142
Db 904 TGTGAGGCTGTGCTGTGAGGACAGACACACCGTGTAAACCGGGGTGTGTATGAT 963
QY 1143 CAGTACTCGGCCACCGAGAGTGTAAATGCAACACCGGCTTCAATGGGACCGGTGTGAG 1202
Db 964 CAGTACAAACCAACAGACAGTGCCTCAATGTCTATCTGTTTCAACGGGACAGCTGTGAG 1023
QY 1203 ATGTGCTGGCCGGGAGATTCGGGCTGTATTTCTGCTGCTGTGCTGTGCTGCTGACACCGA 1262
Db 1024 CTCTGCTGTGCTGGAGATTTGGTCTGACTGTCAACCTGTGGCTGTCTGAGCATGG 1083
QY 1263 CAGTGCAGATGATGGATCAGGGCTCCGGGAGTGCCTGTGTGAAACGGGGTGGACAGGC 1322
Db 1084 CAGTGTGACGAGGGGATCAGAGGCTCCGGGAGTGCCTCTGTGAAGCAGGGTGGACAGGC 1143
QY 1323 CCCTGCTGTGACACTCAGGCAAGTTTGGCTGCAGTGTGTACGCTTCTGCTGTCTGCTCAT 1382
Db 1144 GCCTTCTGTGATGCTCCACAGTGTGTGATCCAGTGTGCATFACCTGTGTTCATGTCAC 1203
QY 1383 GCCACCTGTAAGGAGAAACAACAGCTGTGAGTGTAACTGGATTATGAAGGTGACGGAATC 1442
Db 1204 GCAACCTGTATGGAGAACACAACATGTGTGTAACTTGAATACGAAGGTGACGGAATC 1263
QY 1443 ACATGCAAGTGTGGATTTCTGAAACAGAGACAAACGGGGCTGTGAAAGGTGGCCAGA 1502
Db 1364 ACATGCAAGTGTGGATTTCTGAAACAGAGAACAAATGGGGCTGTGCGAAGGTGGCCAG 1323
QY 1503 TGCTCCCAAGAGGACCAAGGCTCTCCAGAGTCCAGAGGAGGATACAAAGGGAGCGG 1562
Db 1324 TGCTCCCAAGAGGACCAAGTCTCCAGAGTCCAGAGGAGGCTACAAAGGGGATGGC 1383
QY 1563 CACAGCTGCACAGAGATAGACCCCTGTGCAAGACGCGCTTAAACGGAGGGTGTCAAGAGCAC 1622
Db 1384 CACAGCTGCACAGAGATAGACCCCTGTGCAACCGTGTCAATGTGTGATGTCAATGAGCAT 1443
QY 1623 GCCACCTGTAAGATGACAGGCCCGGGGCAAGCACAGTGTGAGTGTAAAGTCACTATGTC 1682
Db 1444 GCTACCTGCAAGATGACGGGGCCAGGCAAGAGTGTGAATGTGAATGTAAAGTCACTATGTC 1503
QY 1683 GGAGATGGGTGAATGTGAGCCGAGCAGCTGCCATTTGACCGCTGCTTACAGACAAAT 1742
Db 1504 GGGATGTGACGGGACTGTGAGCCCGAACAGCTGCCCCCTTGACCGCTGTCTACAGGCAAC 1563
QY 1743 GGGCAGTGCATGACAGAGCCCAATGTGTGACCTCCACTTCCAGGATACCACTGTTGGG 1802
Db 1564 GGACAAATGCCACACAGATGCCAACTGTGTGACCTTCATTTCCAGGACACTACTGTGGC 1623
QY 1803 GTGTTCCATCTACGCTCCCACTGGGCGAGTATTAAGCTGACCTTTGACAAAGCCAGAGAG 1862
Db 1624 GTATTCCACCTTACGCTCCCACTGGGCGAGTACAAACTGACATTTGACAAAGCCAAAGAG 1683
QY 1863 GCCTGTGCAACAGAGTGTGCAATGCAACCTTACACAGGCTCTCTATGCCCCAGAG 1922
Db 1684 GCCTGTGCAACAGAGTGTGCAATGCAACCTTACACAGGCTCTCTATGCCCCAGAG 1743
QY 1923 GCCAAGTACCACTGTGTGCTCAGCAGGCTGGCTGGAGACCGGGGGTGTGCTACCCCA 1982
Db 1744 GCCAAGTATCACCTCTGCTCAGCTGGCTGGCTGGAGAGTGGCGGGTGTGCTACCCCAAG 1803

QY 1983 GCCTCGGCTCCCAAGACTGTGGTCTGTGGTGTGGATAGTGAAGTATGCACTAG 2042
Db 1804 ATCTACGCTCTTAAGAAATGTG---CAACAATTTGGGGATTTGATGATCTACGGAACAGG 1860
QY 2043 CCCAAACAAGAGTGAATGTGGATGTCTTCTGCTATCGGATGAAAGATGTGAAGTGCACC 2102
Db 1861 ACCAAACAAGAGTGAATGTGGGACGCTCTTCTGTACAGAATGAAAGATGTTAACTGCACC 1920
QY 2103 TGCAAGGTGGGTATGTGGGAGATGGCTTCTATGACAGTGGGAACTGCTGAGGTCTCG 2162
Db 1921 TGCAAGGCGAGCTAGCTGGGCGATGGCTTCTCTCAATGGGAACTGCTGCAAGTCTCTC 1980
QY 2163 ATGTCTTTCCCTCACTCACAACTTCTGACGGAAGTGTGGGCTATTTCCAAACAGCTCA 2222
Db 1981 ATGTCTTTCCCTCACTCACGNACTTCTCTGACAGAGTGTGCTGTTTTCAGAGAGCTCA 2040
QY 2223 GCTGAGGCGGTGCATTTCTTAGACAACCTGACTGACTGTCCATCCGGGCAACCTCTTT 2282
Db 2041 GCCCAAGCGCGGGGTTTTGAACACCTGACTGACTGTCCATCAGTGGCAACCTGTGT 2100
QY 2283 GTGCCACAGAACAGTGGGCTGGGGGAGAAATGAGACCTTGTCTGGGCGGAGATCGAGCAC 2342
Db 2101 GTGCCACAGAACAGTGGGCTACCGAATAATAGAGCCTGTCTGGGCGGGAATTGAGCAC 2160
QY 2343 CACCTGCGCAATGTGAGCATGTTTTTCTCAATGACCTTGTCAATGGCACCACTCGCAA 2402
Db 2161 CACCTCACTAATGTCAACGTCTCTCTTTTACGATGACCTTGTCAATGGTACCGTCTGAG 2220
QY 2403 ACGAGGTGGGAAAGCAGCTGCTCATCTGCGAGCGAGACCCACTCCAAACCGAGAG 2462
Db 2221 ACTAGGTGGGAAAGCAGCTGCTCATCACTCCAGCCAGGACCACTCCACCA---AGAG 2277
QY 2463 ACCAGGTGTGTGATGAAGAGCAATCTGCAAGTGGGACATCTTGTGCTTCCAAATGGGATC 2522
Db 2278 GCGAGGTGTGTGATGAAGAGCAATCTGCAAGTGGGACATCTTGTGCTTCAAGGGGT 2337
QY 2523 ATTCATGCTAATTTCCAGGCTTTAAAGACACCCCTGCCCCCGTGACCTTGACCCACACT 2582
Db 2338 CTCCATATCAATTTCTGAACCTTTGAAAGCTCTCCACGGCGGCAACGGCTGCCACTCT 2397
QY 2583 GGCTTGGAGCAGGATCTTCTTTGCCATCATCTCTGTGATCTGGGCTGTGTGCTTGGCT 2642
Db 2398 GGCTTAGGAAACAGGATATCTGTGGCTGTCTGCTGTGTTACTGTGTGAGTGTCTGCT 2457
QY 2643 GCTTACTCTTACTTTCCGATAAACCGGAGAACATCGGCTTCCAGACATTTTGTGTCGAA 2702
Db 2458 GCCTTACTCTTACTTCCGGCTAAACAGAGAACAACTGCTTCCGGCGTTTTGTGATCAGAA 2517
QY 2703 GAGGACATTAATGTGAGCTCTTTGGCAAGCAGGCTGAGAAATATCTCGAACCCCTTG 2762
Db 2518 GATGACATGATGCTTTGGCTTTGGCAAGCAGCAGGCTGAGAGTATCACAAACCTTTG 2577
QY 2763 TATGAGACCAACCTCAGCTCCCCAGAACTTCTTACGACCCCTTACGAGACTCTGAA 2822
Db 2578 TATGAGACCTCAACGCGCGAGCCCCGAGCCCTCTGTGACCCCTTACCGACTCTGGA 2637
QY 2823 GAACGGCAGCTTGAGGCGCAATGACCCCTTGGAGGACACTGTGAGGCGCTTGACGGGAGATG 2882
Db 2638 GAACGGGAGCTGGAGAACAGTACCCCTCTTTGGGCACTG---AGGTCTGTATGAGAG 2694
QY 2883 CCAGCCATCACTCACTGCCACCTGGGCGCATCA-----ACTGTGAATTTCTCA 2928
Db 2695 CCAGCCAAACAGCAGCAGCCAGCTGAGCCCTGAGCCATCAGGTTACACGCTGACTCCCT 2754
QY 2929 GCACAGTTCCTTTAG---GAACGTAAAGTCTTTTAAAGTATGATAAATCTCAGAGCCATACCTCA 2985
Db 2755 GCTCCAGCTGCTCCTTGGATCATTTGTTTTTAAAGTATGATAAATCTCAGAGCCATACCTCA 2814
QY 2986 TCTCTCTGGCTGATCTGGGGGTGTTTCTGTGTGGTGGAGAGATGTGTGCTGCTGCCACCC 3045
Db 2815 CCCTCTGTGTAGTCTGGGTTGTGCTATGGGTAAAGGGACATGTTTCCAGGACACAG 2874

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 The following repetitive elements were found in this cDNA
 sequence: 22-83, >LINE2 (matched complement)
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source
 1..605
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-CO0-ark-g-11-0-UI"
 /tissue_type="mixed"
 /dev_stage="mixed"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Sub9"
 /note="Vector: pTVT3-Pac (Pharmacia) with a modified
 polylinker; Site 1: EcoR I; Site 2: Not I; tissues:
 Cholonic mucosa with Crohns disease, Cholonic mucosa with
 ulcerative colitis, Fetal thymus, Cervix, Cervical
 adenosquamous carcinoma, Ligament cells, Prostate
 carcinoma, Bladder carcinoma, Brain oligodendrocyte ;
 NCI CGAP Sub9 is a subtracted cDNA library constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pTVT3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tags for this library are CGTC, AAGC, GGGCC,
 GAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional
 information, contact: Bento Soares, bento-soares@uiowa.edu
 TAG TISSUE=colonic mucosa with Crohns disease
 TAG LIB=UI-H-CO0
 TAG_SEQ=CGTC"

ORIGIN

Query Match 18.1%; Score 591; DB 14; Length 605;
 Best Local Similarity 99.3%; Pred. No. 6.3e-98;
 Matches 602; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 2653 ACTTTCGATTAACCGGAGAACATCGCGCTCCAGCATTTGAGTCGGAAGAGACATTA 2712
 605 ACTTTCGATTAACCGGAGAACATCGCGCTCCAGCATTTGAGTCGGAAGAGACATTA 546
 2713 ATGTTGACGCTCTTGGCAAGCAGCAGCTGAGATATCTCGAACCCCTTGATGAGACA 2772
 545 ATGTTGACGCTCTTGGCAAGCAGCAGCTGAGATATCTCGAACCCCTTGATGAGACA 486
 2773 CAACCTAGCTCCCGCAGAACCTTCTACGACCCCTTCACGAGCTCTCGAAGACGGCAGC 2832
 485 CAACCTAGCTCCCGCAGAACCTTCTACGACCCCTTCACGAGCTCTCGAAGACGGCAGC 426
 2833 TTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGAGCGGAGATGCCAGCATCA 2892
 425 TTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCCTGAGCGGAGATGCCAGCATCA 366
 2893 CTCACTGCACCTGGGCCATCACTGTGAATTTCTCAGCACCATGTCCTTTAGGACGT 2952
 365 CTCACTGCACCTGGGCCATCACTGTGAATTTCTCAGCACCATGTCCTTTAGGACGT 306
 2953 AAAGTCCTTTAAGCACTCAGAAAGCCATACCTCATCTCTCGGCTGATCTGGGGTGTGTTT 3012
 305 AAAGTCCTTTAAGCACTCAGAAAGCCATACCTCATCTCTCGGCTGATCTGGGGTGTGTTT 246
 3013 CTGTGGGTGAGAGATGTTGCTGTGCCACCCAGTACAGCTTCTCTCTGACCCCTTG 3072
 245 CTGTGGGTGAGAGATGTTGCTGTGCCACCCAGTACAGCTTCTCTCTGACCCCTTG 186
 3073 GCTCTTCTCTTGTACTCTTCAGCTGGCACCTGCTCCTCATCTCTGCCCTAGATGGGT 3132

Db 185 GCTCTTCTCTTGTACTCTTCAGCTGGCACCTGCTCCTCATCTCTGCCCTAGATGGGT 126
 QY 3133 AACTGTGATCTTCTTCCCTGTTAGATGTAAGCCTCCNTCTTTGTATCCAGCCCTAG 3192
 Db 125 AACTGTGA-CTTCTTCCCTGTTAGATGTAAGCCTCCGTCCTTTGTATCCAGCCCTAG 67
 QY 3193 CCAGTGCCTGACACAGGAACCTGTGCACAATAAAGGTTTATGGAACAGAAAAA 3252
 Db 66 CCAGTGCCTGACACAGGAACCTGTGCACAATAAAGGTTTATGGAACAGAAAAA 7
 QY 3253 AAAAAA 3258
 Db 6 AAAAAA 1

RESULT 4

EX101101
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 EX101101
 BX101101 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
 IMAGE998013518 ; IMAGE:241380, mRNA sequence.
 BX101101
 BX101101.1 GI:27831018
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 524)
 Ebert,J., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
 Radelof,U., Schneider,D. and Korn,B.
 Human Unigeneset - RZPD3
 Unpublished (2003)
 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGE998013518.
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
 Human Unigeneset - RZPD3 (RZPDLIB No.972)
 http://www.rzpd.de/CloneCards/cgi-
 bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
 www.rzpd.de
 This clone is available royalty-free from RZPD;
 contact RZPD (clone@rzpd.de) for further information. Seq primer:
 M13r, Primer sequence: TTTCACACAGGAACAGCTATGAC.
 Location/Qualifiers
 1..524
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE998013518 ; IMAGE:241380"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares fetal liver spleen INFLS"
 /note="Organ: Liver and Spleen; Vector: pTVT3D (Pharmacia)
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 [5', AACTGGAAGAAATTAATAAGATCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pTVT3 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M.Fatima Bonaldo."

FEATURES

source
 1..524
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE998013518 ; IMAGE:241380"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares fetal liver spleen INFLS"
 /note="Organ: Liver and Spleen; Vector: pTVT3D (Pharmacia)
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 [5', AACTGGAAGAAATTAATAAGATCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pTVT3 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 15.6%; Score 510; DB 13; Length 524;
 Best Local Similarity 99.2%; Pred. No. 4.3e-83;
 Matches 523; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

```
2729 CAAGCAGCAGCTGAGATATCTCGAACCCCTTGTATGAGAGCACAACCTAGCTCCCCC 2788
Db 1 CAAGCAGCAGCTGAGATATCTCGAACCCCTTGTATGAGAGCACAACCTAGCTCCCCC 60
QY 2789 AGAAGCTTCTAGCAGCCCTTCAAGGACTCTCAAGAGCGGAGCTTGAGGGCAATGACCC 2848
Db 61 AGAAGCTTCTCT---ACCCCTTCAAGGACTCTCAAGAGCGGAGCTTGAGGGCAATGACCC 117
QY 2849 CTTGAGGACACTGTGAGGGCTTGAGCGGAGATGCCAGCATCACTCACTGCCACCTGGG 2908
Db 118 CTTGAGGACACTGTGAGGGCTTGAGCGGAGATGCCAGCATCACTCACTGCCACCTGGG 177
QY 2909 CCATCACTGTGAATCTCAGACCAAGTTCCTTTAGAAAGTAAAGTCTTTAAGCAC 2968
Db 178 CCATCACTGTGAATCTCAGACCAAGTTCCTTTAGAAAGTAAAGTCTTTAAGCAC 237
QY 2969 TCAGAGCCCATACCTCACTCTCTGCTGATCTGGGGTGTCTCTGTGGGTGAGAGATG 3028
Db 238 TCAGAGCCCATACCTCACTCTCTGCTGATCTGGGGTGTCTCTGTGGGTGAGAGATG 297
QY 3029 TGTGCTGTGCCCCACCTAGCAGTTCCTCTGACCCCTTGGCTCTTCTTCTCTTCT 3088
Db 298 TGTGCTGTGCCCCACCTAGCAGTTCCTCTGACCCCTTGGCTCTTCTTCTCTTCT 357
QY 3089 ACTCTTCACTGCGCAGCTGCTCACTCTGCTGATGATGATGATGATGATGATGATGAT 3148
Db 358 ACTCTTCACTGCGCAGCTGCTCACTCTGCTGATGATGATGATGATGATGATGATGAT 417
QY 3149 CCTGTGTAGATTAAGCTTCCTCTGATCCAGCCCTAGCCAGCTGCTGACACA 3208
Db 418 CCTGTGTAGATTAAGCTTCCTCTGATCCAGCCCTAGCCAGCTGCTGACACA 477
QY 3209 GGAACCTGTGCACATAAGGTTTATGGAACAGAAAAA 3255
Db 478 GGAACCTGTGCACATAAGGTTTATGGAACAGAAAAA 524

RESULT 5
BG220572
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

BG220572 562 bp mRNA linear EST 21-APR-2001
RST40358 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
BG220572
BG220572.1 GI:13746593
EST.
Homo sapiens
Homo sapiens (human)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Boozer, S.,
Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,
Offenbacher, J., Danzig, J. and Ducar, M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
21227151
11329013
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 399.
Location/Qualifiers
1..562
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
```

ORIGIN

```
Query Match
Best Local Similarity 14.9%; Score 486; DB 12; Length 562;
Matches 534; Conservative 0; Mismatches 27; Indels 3; Gaps 3;
QY 2690 TTTTCAGTCGGAAGAGGACATTAATGTTGACGCTCTTGCGAAGCAGCAGCCTTGAGAATAT 2749
Db 2 TTGTGTGGGACTGAGGACATTAATGTTGACGCTCTTACCAGCAGTACCTGAGATAT 61
QY 2750 CTCGAACCCCTTGTATGAGAGCAACCTCAGCTCCCCAGAACCTTCTCAGACCCCTT 2809
Db 62 CTCGAATCCCTTGTATGAGAGCAACCTCAGCTCCCCAGAACCTTCTCAGACCCCTT 121
QY 2810 CACGAGCTCTGAAGAACGCGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTAGGGCC 2869
Db 122 CACGAGCTCTGAAGAACGCGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTAGGGCC 181
QY 2870 TGGACGGGAGATGCCAGGCCATCACTCAGCTGCCACCTGGGGCCATCAACTGTGAATCTCAG 2929
Db 182 TGGACGGGAGATGCCAGGCCATCACTCAGCTGCCACCTGGGGCCATCAACTGTGAATCTCAG 241
QY 2930 CACCAAGTTCCTTTTAGGAAACGTAAGTCTTTAAGCACTCAGAGCCCATCTCATCTC 2989
Db 242 CACCAAGTTCCTTTTAGGAAACGTAAGTCTTTAAGCACTCAGAGCCCATCTCATCTC 301
QY 2990 TCTGGCTGATCGGGGTTTCTCTGAGGTCAGAGATGTTGCTGTGCCACCCAGTA 3049
Db 302 TCTGGCTGATCGGGGTTTCTCTGAGGTCAGAGATGTTGCTGTGCCACCCAGTA 361
QY 3050 CAGCTTCCTCTCTGACCCCTTTGGCTCTTCTCTTTGTACTCTTCAGCTGGCACCTGCT 3109
Db 362 CAGCTTCCTCTCTGACCCCTTTGGCTCTTCTCTTTGTACTCTTCAGCTGGCACCTGCT 421
QY 3110 CCATTCTGCCCTACATGATGGGTAATGATGATCTTCTTCCCTGTTAGATTTGAAGCTC 3169
Db 422 CCATTCTGCCCTACATGATGGGTAATGATGATCTTCTTCCCTGTTAGATTTGAAGCTC 479
QY 3170 CMTCTTTGATCCAGCCCTAGCCAGTCCCTGACACAGGACTCTGCACATAAAGT 3229
Db 480 CAGCTTTGTATCCCA-CCCTTAACCAATGCTGACACAGGACTGTGCACACAAAAGG 538
QY 3230 TTATGGAACAGAAAAA 3253
Db 539 TTATGGAACAGAAAAA 562
```

RESULT 6

```
AI081595/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AI081595 481 bp mRNA linear EST 28-AUG-1998
ox63h05.s1 Soares_NhMPu_S1 Homo sapiens cDNA clone IMAGE:1661049
3'', mRNA sequence.
AI081595
AI081595.1 GI:3418387
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 481)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncilogap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
```


Db 84 GCCTCGTCTTTGTATCCAGCCCTAGCCAGTCTGTACACAGAGACTGTGTCAATA 25

QY 3225 AAGTTTATGGAACAGAAAAAAA 3248
 |||||
 Db 24 AAGTTTATGGAACAGAAAAAAA 1
 |||||

RESULT 8
 A1131339/c
 LOCUS
 DEFINITION q18c08.x1 Soares fetal heart NBHH19W Homo sapiens cDNA clone
 IMAGE:1709966 3', mRNA sequence.

ACCESSION A1131339
 VERSION A1131339.1 GI:3601355
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 421)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (infoimage.llnl.gov) for further information.
 Insert Length: 364 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham.
 Location/Qualifiers
 1..421
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:1709966"
 /sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares fetal heart NBHH19W"
 /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5',
 TGTACCAATCTGAAGTGGAGCGCGGCACATCTTTTTTTTTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by
 M.Fatima Bonaldo. This library was constructed from the
 same fetus as the fetal lung library, Soares fetal lung
 NBHL19W."

ORIGIN

Query Match 12.8%; Score 418.4; DB 9; Length 421;
 Best Local Similarity 99.5%; Pred. No. 2.5e-66;
 Matches 419; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2824 AACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGCGCTGGACGGAGATGC 2883
 Db 421 AACGGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGCGCTGGACGGAGATGC 362

QY 2884 CAGCCATCATCTCACTGCGCACCTGGGCCCATCAACTGTGAAATTCTCAGCACAGTTGCCCTT 2943
 Db 361 CAGCCATCATCTCACTGCGCACCTGGGCCCATCAACTGTGAAATTCTCAGCACAGTTGCCCTT 302

QY 2944 TAGGAACGTAAGTCCTTTAAGCACTCAGAGCCATACCTCATCTCTTGCTGCTATCTCG 3003
 Db 301 TAGGAACGTAAGTCCTTTAAGCACTCAGAGCCATACCTCATCTCTTGCTGCTATCTCG 242

QY 3004 GGGTTGTTCTTGCGGTGAGAGATGTGTTGCTGTGCCACCCAGTACAGTTCCTCTCT 3063
 Db 241 GGGTTGTTCTTGCGGTGAGAGATGTGTTGCTGTGCCACCCAGTACAGTTCCTCTCT 182

QY	2843	TAACCCCTT	AGGACAC	TGTGAGGGCC	TGGAGGAGATG	CCAGCCATCACT	CACTGCGCA	2902
Db	403	TGACCCCTT	TGAGGAACT	TGTGAGGNC	CTTGGACGGGAGATG	CNAGCCATCACT	CACTGCGCA	344
QY	2903	CCTGGG-CA	CACTGTAAT	TCACAGA-CC	AGTGTGCTTT	TAGGAACGTAA	AGTCTCT	2960
Db	343	CCCGGGCC	CACTCACT	GTGAATTC	CAGACCCAGT	TGCTTTTAGGA	ACGTAAAGTCT	284
QY	2961	TTAAGCACT	CAGAGCCAT	ACCTCATCT	CTCTGGCTGAT	CTGGGGGTTG	TTTCTGTGGT	3020
Db	283	TTAAGCACT	CAGAGCCAT	ACCTCATCT	CTCTGGCTGAT	CTGGGGGTTG	TTTCTGTGGT	224
QY	3021	GAGAGATG	TGTGTGCT	GGCCACCC	AGTACAGTCT	CTCTCTGAC	CCCTTGGCTCTTCT	3080
Db	223	GAGAGATG	TGTGTGCT	GGCCACCC	AGTACAGTCT	CTCTCTGAC	CCCTTGGCTCTTCT	164
QY	3081	TCCTTTGAC	CTCTCAG	CTGGCACCT	GTCTCCATT	CTGCCCTAC	GATGGGTAAC	3140
Db	163	TCCTTTGAC	CTCTCAG	CTGGCACCT	GTCTCCATT	CTGCCCTAC	GATGGGTAAC	104
QY	3141	TCCTTTTCC	TCTGTTAG	ATGTAAGCT	CTCCNTCTT	TGTAATCC	AGCCCTAGCCAGTGC	3200
Db	103	TCCTTTTCC	TCTGTTAG	ATGTAAGCT	CTCCNTCTT	TGTAATCC	AGCCCTAGCCAGTGC	44
QY	3201	CTGACACAG	GAAGTGTG	CAATAAAG	GTATGGAACA	GAGA	3241	
Db	43	CTGACACNG	GAAGTGTG	CAATAAAG	GTATGGAANGA	3		
RESULT 11								
AA968560/c								
LOCUS								
DEFINITION								
ACCESSION								
VERSION								
KEYWORDS								
SOURCE								
ORGANISM								
REFERENCE								
AUTHORS								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
source								

Seq primer: M13RP1
High quality sequence stop: 362.
Location/Qualifiers
1. .448
source

FEATURES
source

11.7%; Score 381; DB 14; Length 448;
Best Local Similarity 99.3%; Pred. No. 1.8e-59;
Matches 403; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 2728 GCAAGCAGCAGCTGAGATATCTCGAACCCCTTGTATGAGAGCACAACCTCAGCTCCCC 2787
Db 1 GCAAGCAGCAGCTGAGATATCTCGAACCCCTTGTATGAGAGCACAACCTCAGCTCCCC 60

QY 2788 CAGAACCTTCTAGACCCCTTACGAGCTCTGAGAGCAGGAGCTTGGGCAATGACC 2847
Db 61 CAGAACCTTCTTACGACCCCTTACGAGCTCTGAGAGCAGGAGCTTGGGCAATGACC 120

QY 2848 CTTGAGGACACTGTGAGGGCTGAGCGGAGATCCAGCCATCACTCACTGCCACCTGG 2907
Db 121 CTTGAGGACACTGTGAGGGCTGAGCGGAGATCCAGCCATCACTCACTGCCACCTGG 180

QY 2908 GCCATCACTGTGAATCTCAGCACCAGTTCCTTTAGGAACGTAAGTCTCTTAAAGCA 2967
Db 181 GCCATCACTGTGAATCTCAGCACCAGTTCCTTTAGGAACGTAAGTCTCTTAAAGCA 240

QY 2968 CTCAGAGCCATACCTATCTCTCTGGTGATCTGGGGTTTCTGTGGGTGAGAGAT 3027
Db 241 CTCAGAGCCATACCTATCTCTCTGGTGATCTGGGGTTTCTGTGGGTGAGAGAT 300

QY 3028 GTGTGCTGTGCCACCCAGTACAGCTTCCTCTGACCCCTTGGCTCTTCTTCTTTG 3087
Db 301 GTGTGCTGTGCCACCCAGTACAGCTTCCTCTGACCCCTTGGCTCTTCTTCTTTG 360

QY 3088 TACTCTTCAGCT-GGCACCTGTCTCAATTC-TGCCCTACATGATGGG 3131
Db 361 TACTCTTCAGCTGGGACCTGCTCCATTNTTGGCCCTACATGATGGG 406

RESULT 13
AA010567/c
LOCUS
DEFINITION
484 bp mRNA linear EST 09-MAY-1997
clone IMAGE:430269 3', mRNA sequence.
ACCESSION
AA010567
VERSION
AA010567.1 GI:1471613
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 484)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., Dubucque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,

Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
PUBMED
8889549
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewartson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 946 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 307.
Location/Qualifiers
1. .484
source

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1330038"
/db_xref="taxon:9606"
/clone="IMAGE:430269"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
This is a subcloned version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I oligo (dT) primer [5',
AAGTGAAGAAATTAATAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

QY 2821 AAGAACGCGAGCTTGAGGCAATGACCCCTTGAGGACACCTGAGGGCTTGACGGGAGA 2880
Db 425 AAGAACGCGAGCTTGAGGCAATGACCCCTTGAGGACACCTGAGGGCTTGACGGGAGA 366

QY 2881 TGCCAGCCATCACTCAGTCCCA--CCTGGGCCATCACTGTAATTCAGCACC-AGTT 2937
Db 365 TGCCAGCCATCACTCAGTCCCACTGGGCCATCACTGTAATTCAGCACCAGTT 306

QY 2938 GCCTTTTAGGAACGTAAGTCTTTAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTG 2997
Db 305 GCCTTTTAGGAACGTAAGTCTTTAAGCACTCAGAAGCCATACCTCATCTCTGGCTG 246

QY 2998 ATCTGGGGTGTCTTCTGTGGGTGAGAGATGTGTGTCGCCACCCAGTACAGTTCC 3057
Db 245 ATCTGGGGTGTCTTCTGTGGGTGAGAGATGTGTGTCGCCACCCAGTACAGTTCC 186

QY 3058 TCCTCTGACCCCTTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3116
Db 185 TCCTCTGACCCCTTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 126

QY 3117 GCCCTACATGATGGTAACTGTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3176
Db 125 GCCCTACATGATGGTAACTGTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 66

QY 3177 GTATCCAGCCCTTAGCCAGTGTGCTGACAGAACTGTGACAAATAAGGTTTATGGA 3236
Db 3177 GTATCCAGCCCTTAGCCAGTGTGCTGACAGAACTGTGACAAATAAGGTTTATGGA 3236

```

Db      65  GTATCCAGCCCTAGCCAGTGCCTGACACAGGAAGTGTGCACAATTAAGCTTTATGGA 6
QY      3237 ACAGA 3241
Db      5  ACAGA 1

RESULT 14
BC049247
LOCUS      3025 bp      mRNA      linear      HTC 28-MAR-2003
DEFINITION Mus musculus, Similar to stabilin 1, clone IMAGE:5038209, mRNA.
ACCESSION BC049247
VERSION   BC049247.1  GI:29351595
KEYWORDS  HTC.
SOURCE    Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS   Strausberg,R.
TITLE     Direct Submission
JOURNAL   Submitted (21-MAR-2003) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
          Email: cgabs.rem@nih.gov
          Tissue Procurement: Gilbert Smith, Ph.D.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NIHSC),
          Gaithersburg, Maryland;
          Web site: http://www.nisic.nih.gov/
          Contact: nisc.mgc@nih.gov
          Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
          Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
          Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
          Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Lalic,P., Legaspi,R.,
          Madowo,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
          McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
          Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
          Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 99 Row: g Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source            1..3025
     organism="Mus musculus"
     mol_type="mRNA"
     strain="CZECH II"
     db_xref="taxon:10090"
     clone="IMAGE:5038209"
     tissue types="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clone) outgrowth infected with the virus MMTV."
     clone_1ib="NCI CGAP_Lu29"
     lab_host="DH10B"
     note="vector: pCMV-SPORT6"

ORIGIN
Query Match      11.4%; Score 370.4; DB 11; Length 3025;
Best Local Similarity 54.8%; Pred. No. 1.1e-57;
Matches 828; Conservative 0; Mismatches 666; Indels 18; Gaps 4;
QY      1046 GATCCCGAGTGTGCGAGGCTACTTCGGCGGAGACTGTGAGCCCTGCCCTGGAGACC 1105
          || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2004, 10:14:56 ; Search time 1177 Seconds
(without alignments)
11766.462 Million cell updates/sec

Title: US-10-028-248A-210
Perfect score: 3260
Sequence: 1 95cagcagcagagcttcccc.....aaaaaaaaaaaaaaaaaaaaa 3260

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2003as:*
- 8: geneseqn2003bs:*
- 9: geneseqn2003cs:*
- 10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3259	100.0	3260	ABT08489	Abt08489 Human nov
2	3232.6	99.2	4576	ABA04662	AbA04662 Human Hya
3	3232.6	99.2	4642	ABX13822	Abx13822 cDNA enco
4	2999.8	92.0	8444	ABT08487	Abt08487 Human nov
5	2995.6	91.9	8495	ABT08488	Abt08488 Human nov
6	2853.2	87.5	4962	ABX13823	Abx13823 cDNA enco
7	2828.8	86.8	3625	AAf87120	AAf87120 NOV9 codi
8	1936.8	59.4	4706	ABA04648	AbA04648 Rat Hyalu
9	1936.8	59.4	4706	ABX13821	Abx13821 cDNA enco
10	1778.6	54.6	2483	AAf87119	AAf87119 NOV8 codi
11	1777	54.5	2512	AAf87117	AAf87117 NOV6 codi
12	1721.2	52.8	2011	AAf87113	AAf87113 NOV2 codi
13	1429.6	43.9	1804	AAf87114	AAf87114 NOV3 codi
14	1347.4	41.3	1377	AAf87114	AAf87114 NOV3 codi
15	1227.4	37.7	1259	AAA57365	AAa57365 DNA enco
16	990.2	30.4	1624	AAf87118	AAf87118 NOV7 codi
17	764.2	23.4	1577	AAf87121	AAf87121 NOV10 codi
18	708	21.7	718	AAf87121	AAf87121 NOV10 codi
19	537	16.5	722	AAf87121	AAf87121 NOV10 codi
20	424.6	13.0	3681	AAf87121	AAf87121 NOV10 codi
21	423.4	13.0	7879	AAf87121	AAf87121 NOV10 codi
22	421.8	12.9	6761	AAf87121	AAf87121 NOV10 codi
23	408.8	12.5	448	AAf87121	AAf87121 NOV10 codi

24	391.2	12.0	2863	9	ADB62456	ADB62456 Human cdn
25	285	8.7	2160	9	ADB63670	ADB63670 Human cdn
26	259.6	8.0	2753	9	ADB47730	ADB47730 Novel hum
27	248	7.6	1522	3	AAa57363	AAa57363 DNA enco
28	247.4	7.6	395	8	ACH29524	ACH29524 Human adu
29	187.4	5.7	1482	7	ACH55632	ACH55632 Human sig
30	143	4.4	473	8	ACH15198	ACH15198 Human adu
31	135	4.1	420	8	ACH21247	ACH21247 Human adu
32	125.2	3.8	474	8	ACH45565	ACH45565 Human foe
33	81.4	2.5	781	7	ABX74436	ABx74436 Human cdn
34	81.4	2.5	1144	7	ABX63446	ABx63446 Human cdn
35	81.4	2.5	1144	9	ABE25720	ABe25720 Human cdn
36	81.4	2.5	1414	2	AAQ27190	AAq27190 Tumour ne
37	81.4	2.5	1414	2	AAV71778	AAv71778 Tumour ne
38	81.4	2.5	1414	4	AAQ06019	AAQ06019 Human tum
39	81.4	2.5	1414	6	ABs54635	ABs54635 Human cdn
40	81.4	2.5	1422	7	ACC72678	ACC72678 Human can
41	81.4	2.5	1422	7	ABX76382	ABx76382 Lung canc
42	81.4	2.5	1728	7	ABX63445	ABx63445 Human cdn
43	78.2	2.4	351	6	ABV95472	ABv95472 Human pan
44	78.2	2.4	1414	6	ABL68341	ABl68341 Kidney ca
45	78.2	2.4	1414	6	ABK83989	ABk83989 Human cdn

ALIGNMENTS

RESULT 1

ABT08489

ID ABT08489 standard; cDNA; 3260 BP.

XX AC

XX ABT08489;

XX DT

XX 28-NOV-2002 (first entry)

XX DE

XX Human novel protein NOV1c coding sequence SEQ ID NO: 210.

XX KW

Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
antidepressant; immunosuppressive; antibacterial; antiparasitic;
virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
antiaddictive; gene; ss.

XX OS

XX Homo sapiens.

XX EH

XX Key Location/Qualifiers

XX FT

XX variation replace(887,C)

XX FT

XX variation /*tag= a

XX FT

XX variation replace(1034,T)

XX FT

XX variation /*tag= b

XX XX

XX variation /*tag= c

XX PN

XX WO200259315-A2.

XX PD

XX 01-AUG-2002.

XX XX

XX 19-DEC-2001; 2001WO-US050076.

XX PF

XX 19-DEC-2000; 2000US-0256619P.

XX PR

XX 19-JAN-2001; 2001US-0262959P.

XX PR

XX 28-FEB-2001; 2001US-0272408P.

XX PR

XX 26-APR-2001; 2001US-0285189P.

XX PR

XX 20-JUL-2001; 2001US-0308039P.

XX PR

XX 09-AUG-2001; 2001US-0311266P.

XX XX

XX (CURA-) CURAGEN CORP.

XX PA

Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;
Shenoy S, Spytek RA, Gargolli E, Miller C, Boldog F, Li L;
Taupier Ru, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;


```
Db 1691 TCGGAGATGGGCTGAACCTGTGAGCGGAGCAGCTGCCATTGACCGCTGTACAGGACA 1740
QY 1741 ATGGCAGTGGCCTCAGACGCGCAATATGTGTGAGCTCCACCTTCAGAGTACCACTGTG 1800
Db 1741 ATGGCAGTGGCCTCAGACGCGCAATATGTGTGAGCTCCACCTTCAGAGTACCACTGTG 1800
QY 1801 GGGTGTTCATCTAGCTTCCCTCCCTGAGGCGAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860
Db 1801 GGGTGTTCATCTAGCTTCCCTCCCTGAGGCGAGTATAAGCTGACCTTTGACAAAGCCAGAG 1860
QY 1861 AGGCTGTGCGCAACGAAGCTGGACATGGCAACCTACAAACAGCTCTCCCTATGCCGAGA 1920
Db 1861 AGGCTGTGCGCAACGAAGCTGGACATGGCAACCTACAAACAGCTCTCCCTATGCCGAGA 1920
QY 1921 AGGCCAAGTACCACTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTGTGCTACCCCA 1980
Db 1921 AGGCCAAGTACCACTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTGTGCTACCCCA 1980
QY 1981 CAGCCTTCGCTCCAGAACTGTGGCTCTGTGTGGTGGGATAGTGAACCTATGAGACCTA 2040
Db 1981 CAGCCTTCGCTCCAGAACTGTGGCTCTGTGTGGTGGGATAGTGAACCTATGAGACCTA 2040
QY 2041 GACCCAAAGAGTGAATGTGGGATGCTCTGTCTATCGGATGAAGATGTGAACCTGA 2100
Db 2041 GACCCAAAGAGTGAATGTGGGATGCTCTGTCTATCGGATGAAGATGTGAACCTGA 2100
QY 2101 CCTCAAGTGGGCTATGTGGGAGATGGCTTCTCATGCAAGTGGGAACTGTGCAAGTCC 2160
Db 2101 CCTCAAGTGGGCTATGTGGGAGATGGCTTCTCATGCAAGTGGGAACTGTGCAAGTCC 2160
QY 2161 TGATGTCTTCCCTCACTCAAACTTCTGCAAGGAGTGGCTTCTCATGCAAGTGGGAACTGTGCAAGTCC 2220
Db 2161 TGATGTCTTCCCTCACTCAAACTTCTGCAAGGAGTGGCTTCTCATGCAAGTGGGAACTGTGCAAGTCC 2220
QY 2221 CAGCTCGAGGCGGTGCACTTCTAGAACACCTGACCTGACCTGCTCCATCGCGGACCTCT 2280
Db 2221 CAGCTCGAGGCGGTGCACTTCTAGAACACCTGACCTGACCTGCTCCATCGCGGACCTCT 2280
QY 2281 TTGTGCCACAGACAGTGGGCTGGGGAGATGAGACCTTGTCTGGGGGGGACATCGAGC 2340
Db 2281 TTGTGCCACAGACAGTGGGCTGGGGAGATGAGACCTTGTCTGGGGGGGACATCGAGC 2340
QY 2341 ACCACCTCGCAATGTGAGATGTTTTTCTCAATGACCTTGTCAATGGCAACCACTGTC 2400
Db 2341 ACCACCTCGCAATGTGAGATGTTTTTCTCAATGACCTTGTCAATGGCAACCACTGTC 2400
QY 2401 AAACGAGGCTGGGAGCAAGTGTCTCATCTGCGCAGCAGACCCCTCAACCGAGCG 2460
Db 2401 AAACGAGGCTGGGAGCAAGTGTCTCATCTGCGCAGCAGACCCCTCAACCGAGCG 2460
QY 2461 AGACAGGTTTGTGATGGAAGACCAATCTGAGTGGGACATCTTTGCTCCCAATGGGA 2520
Db 2461 AGACAGGTTTGTGATGGAAGACCAATCTGAGTGGGACATCTTTGCTCCCAATGGGA 2520
QY 2521 TCATTCATGCTATTCAGGCTTTAAAGACACCCCTGCGCGTGAACCTGACCTGACCCACA 2580
Db 2521 TCATTCATGCTATTCAGGCTTTAAAGACACCCCTGCGCGTGAACCTGACCTGACCCACA 2580
QY 2581 CTGGCTTGGGAGCAGGATCTTCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Db 2581 CTGGCTTGGGAGCAGGATCTTCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
QY 2641 CTGCTTACTCTACTTTCGGAATAACCGGAGCAATCGGCTTCAGCATTTTCAGTCGG 2700
Db 2641 CTGCTTACTCTACTTTCGGAATAACCGGAGCAATCGGCTTCAGCATTTTCAGTCGG 2700
QY 2701 AAGAGGACATTAATGTGAGCTCTTGGCAAGCAGCAGCCTGAGAAATATCTCGAACCCCT 2760
Db 2701 AAGAGGACATTAATGTGAGCTCTTGGCAAGCAGCAGCCTGAGAAATATCTCGAACCCCT 2760
QY 2761 TGTATGAGACACACCTCAGCTCCCGCAGAACCTTCTAGACCCCTTACGAGCTCTG 2820
```

```
Db 2761 TGTATGAGACACACCTCAGCTCCCGCAGAACCTTCTTACGACCCCTTACGAGCTCTG 2820
QY 2821 AAGAACGCGCAGCTTGGGGCAATGACCCCTTGGAGACACTGTGAGGCGCTGACGCGGAGA 2880
Db 2821 AAGAACGCGCAGCTTGGGGCAATGACCCCTTGGAGACACTGTGAGGCGCTGACGCGGAGA 2880
QY 2881 TGCCAGCCATCACTCACTGCGCACCTGGGCCATCAACTGTGAATCTCAGCACCACTTGC 2940
Db 2881 TGCCAGCCATCACTCACTGCGCACCTGGGCCATCAACTGTGAATCTCAGCACCACTTGC 2940
QY 2941 TTTTAGGAACCTAAGTCTTTAAGCACTCAGAGCCATACCTCATCTCTCTGGCTGATC 3000
Db 2941 TTTTAGGAACCTAAGTCTTTAAGCACTCAGAGCCATACCTCATCTCTCTGGCTGATC 3000
QY 3001 TGGGGTGTGTTCTGTGGGTGAGAGATGTGTGCTGCTGCCACCCAGTACAGCTTCTCTCC 3060
Db 3001 TGGGGTGTGTTCTGTGGGTGAGAGATGTGTGCTGCTGCCACCCAGTACAGCTTCTCTCC 3060
QY 3061 TCTGACCCCTTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3120
Db 3061 TCTGACCCCTTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3120
QY 3121 TACATGATGGCTAACTGTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3180
Db 3121 TACATGATGGCTAACTGTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3180
QY 3181 CCCAGCCCTAGCCAGTGCCTGACACAGGAACCTGTGCACAATAAAGTTTATGGAACAG 3240
Db 3181 CCCAGCCCTAGCCAGTGCCTGACACAGGAACCTGTGCACAATAAAGTTTATGGAACAG 3240
QY 3241 AAAAAAAAAAAAAAAAAAAAAA 3260
Db 3241 AAAAAAAAAAAAAAAAAAAAAA 3260
```

RESULT 2

ABA04662

ID ABA04662 standard; cDNA; 4576 BP.

XX ABA04662;

XX 22-FEB-2002 (first entry)

XX Human Hyaluronic Acid Receptor for Endocytosis, HARE, coding sequence.

XX HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;

XX chondroitin sulphate; extracellular matrix; cartilage; skin;

XX vitreous humour; endocytic receptor; glycosaminoglycan; human; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..4185

XX /*tag= a

XX /partial

XX /product= "Human HARE"

XX /note= "No start codon given"

XX WO200181544-A2.

XX 01-NOV-2001.

XX 25-APR-2001; 2001WO-US013403.

XX 25-APR-2000; 2000US-0199538P.

XX 02-NOV-2000; 2000US-0245320P.

XX (WEIG/) WEIGEL P A.

XX (ZHOU/) ZHOU B.

XX (WEIG/) WEIGEL J A.

XX Weigel PA, Zhou B, Weigel JA;

XX

DR WPI, 2002-049271/06.
DR P-PSDB; AAM47684.
XX
PT New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for
PT identifying agents that inhibit binding to hyaluronic acid, and related
PT nucleic acid.
XX
PS Claim 49; Fig 33; 263pp; English.
XX
CC The present invention relates to sequences for rat and human HARE
CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675
CC and AAM47684). HARE can bind specifically to at least one of hyaluronic
CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin
CC sulphate (CS). HA is an extracellular matrix component of all tissues, in
CC particular cartilage, skin and vitreous humour. HARE is the endocytic
CC receptor responsible for removing HA and other glycosaminoglycans from
CC the circulation. The present sequence is the coding sequence for human
CC HARE
XX
SQ Sequence 4576 BP; 1127 A; 1193 C; 1203 G; 1053 T; 0 U; 0 Other;
Query Match 99.2%; Score 3232.6; DB 6; Length 4576;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3238; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
7 AGCAGAGCTCCCAAGAACCCGAAACCTTCCAGTATTTCTTCCAGTTTGCAGGAGCAT 66
1328 ATCAGAGCTTCCCAAGAACCCGAAACCTTCCAGTATTTCTTCCAGTTTGCAGGAGCAT 1387
67 TCGTGAAGAGCTCGGTGCGGCCAGGCCCTTCCACTGTTTTTGCACCTTTATCTGAGCCT 126
1388 TCGTGAAGAGCTCGGTGCGGCCAGGCCCTTCCACTGTTTTTGCACCTTTATCTGAGCCT 1447
127 TTGATGAGAGCTCGGTGCGGCCAGGCCCTTCCACTGTTTTTGCACCTTTATCTGAGCCT 186
1448 TTGATGAGAGCTCGGTGCGGCCAGGCCCTTCCACTGTTTTTGCACCTTTATCTGAGCCT 1507
187 GGTACCATCTGTCGCTCCGACCGAGCTGCTTCTGGAACCTGAAATGATCTCAAATG 246
1508 GGTACCATCTGTCGCTCCGACCGAGCTGCTTCTGGAACCTGAAATGATCTCAAATG 1567
247 CTACTTCCCTCCAAAGGAGCAATAGTATCATCTCGTCTCTCAGAGCAGCGTGTATATAA 306
1568 CTACTTCCCTCCAAAGGAGCAATAGTATCATCTCGTCTCTCAGAGCAGCGTGTATATAA 1627
307 ATAAATAGGCTAAGATCATATCAGTGTATCATATGATGATGATGATGATGATGATGAT 366
1628 ACAATAAGGCTAAGATCATATCAGTGTATCATATGATGATGATGATGATGATGATGAT 1687
367 TAGACAAATTTGCTATCTCCAAATTTTCTTATCATCTCCAAAGCAACTCTGGAAGAA 426
1688 TAGACAAATTTGCTATCTCCAAATTTTCTTATCATCTCCAAAGCAACTCTGGAAGAA 1747
427 TTCTGCAAAATCTTACGACTTTGGCAACAAATGGCTGATCATCAATTTAGCACTTAA 486
1748 TTCTGCAAAATCTTACGACTTTGGCAACAAATGGCTGATCATCAATTTAGCACTTAA 1807
487 TACAGACTCAGTTTGTGCTGAGTGTATCATCCGATCCCATCCACCCAGTCACTCTCT 546
1808 TACAGACTCAGTTTGTGCTGAGTGTATCATCCGATCCCATCCACCCAGTCACTCTCT 1867
547 TCTGGCCCAACGACCAAGCCCTCCATGCTTACCTGCTGAAACAAAGGACTTCTCTGTCA 606
1868 TCTGGCCCAACGACCAAGCCCTCCATGCTTACCTGCTGAAACAAAGGACTTCTCTGTCA 1927
607 ACCAAGACAAAGGCAAGCTGAAGAGTATTTGAAGTTTATGATGATGATGATGATGATGAT 666
1928 ACCAAGACAAAGGCAAGCTGAAGAGTATTTGAAGTTTATGATGATGATGATGATGATGAT 1987
667 AGTGTGCTGCTGCTGCTTCCACATCCACTGCTGGAAGCCCTGCAAGGTTTCAAGC 726
1988 AGTGTGCTGCTGCTTCCACATCCACTGCTGGAAGCCCTGCAAGGTTTCAAGC 2047

QY 727 TGAGTGTCAAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 786
DB 2048 TGAGTGTCAAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTTCTGAATGGCCAAACCT 2107
QY 787 GCAGAAATGTGACGGGAGCTCTTGTGTTGACCTGGGTGTGGCCTTACGGCAATGATGTC 846
DB 2108 GCAGAAATGTGACGGGAGCTCTTGTGTTGACCTGGGTGTGGCCTTACGGCAATGATGTC 2167
QY 847 TGCTGATTTGATCCACCTCTGGGGGCGCTGTGCACCTTTACTTACTTTTCCGATGCTCGG 906
DB 2168 TGCTGATTTGATCCACCTCTGGGGGCGCTGTGCACCTTTACTTACTTTTCCGATGCTCGG 2227
QY 907 GGGAGTGTGGAGCTGTGTCAATACTCCACAGTGCCTCCAGGTGAGTAAACAAAGGGTG 966
DB 2228 GGGAGTGTGGAGCTGTGTCAATACTCCACAGTGCCTCCAGGTGAGTAAACAAAGGGTG 2287
QY 967 TGAAGCAGAAGTGTCTTACAACTGCTTCAAGAGGAACTTGAAGGCTCCCGGGAGC 1026
DB 2288 TGAAGCAGAAGTGTCTTACAACTGCTTCAAGAGGAACTTGAAGGCTCCCGGGAGC 2347
QY 1027 GGTGCAGCTGTGTGATACAGATCCCGAGTGTCTCAAGGGCTTACTTCGGGGCGAGACTGTC 1086
DB 2348 GGTGCAGCTGTGTGATACAGATCCCGAGTGTCTCAAGGGCTTACTTCGGGGCGAGACTGTC 2407
QY 1087 AGGCTCTGCCCTGGAGGACAGATGCCCGTGTAAATACCGGGGTGTCTGCTTGTGATCAGT 1146
DB 2408 AGGCTCTGCCCTGGAGGACAGATGCCCGTGTAAATACCGGGGTGTCTGCTTGTGATCAGT 2467
QY 1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGAGCGGTGTGAGATGT 1206
DB 2468 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGAGCGGTGTGAGATGT 2527
QY 1207 GCTGCGCGGGAGATTTCCGGCTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1266
DB 2528 GCTGCGCGGGAGATTTCCGGCTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2587
QY 1267 GCGATGATGGCATCACGGGCTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1326
DB 2588 GCGATGATGGCATCACGGGCTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2647
QY 1327 CGTGTGACACTCAGGAGTGTTCCTGAGTGTGACCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1386
DB 2648 CGTGTGACACTCAGGAGTGTTCCTGAGTGTGACCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2707
QY 1387 CCTGTAAAGAGAACCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
DB 2708 CCTGTAAAGAGAACCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2767
QY 1447 GCACAGTTGTGGATTTCTGCAACAGGACCAACCGGGGCTGTGCAAGGTTGGCCAGATGCT 1506
DB 2768 GCACAGTTGTGGATTTCTGCAACAGGACCAACCGGGGCTGTGCAAGGTTGGCCAGATGCT 2827
QY 1507 CCAGAGGGCAGAGGTTCTCTGAGTGTGACAGAGGATACAAAGGGGAGCGGACCA 1566
DB 2828 CCAGAGGGCAGAGGTTCTCTGAGTGTGACAGAGGATACAAAGGGGAGCGGACCA 2887
QY 1567 GCTGCACAGAGATAGACCCCTGTGACAGCGGCTTAAACGAGGGTGTTCACGAGCAGCCCA 1626
DB 2888 GCTGCACAGAGATAGACCCCTGTGACAGCGGCTTAAACGAGGGTGTTCACGAGCAGCCCA 2947
QY 1627 CCTGTAAAGATGACAGCGCGGCGCAACAGTGTGAGTGTAAAGTGTCTATGTCTGGAG 1686
DB 2948 CCTGTAAAGATGACAGCGCGGCGCAACAGTGTGAGTGTAAAGTGTCTATGTCTGGAG 3007
QY 1687 ATGGGCTGAATGTGAGCGGAGAGCTGCCCATTTGACCGCTGCTTACAGGACAATGGGC 1746
DB 3008 ATGGGCTGAATGTGAGCGGAGAGCTGCCCATTTGACCGCTGCTTACAGGACAATGGGC 3067
QY 1747 AGTGGCATGACAGCGCAAAATGTGTGAGCTTCCAGGATACCACTGTTGGGGTGT 1806
DB 3068 AGTGGCATGACAGCGCAAAATGTGTGAGCTTCCAGGATACCACTGTTGGGGTGT 3127
QY 1807 TCCATCTACGCTCCCACTGGGCGAGTATAAGTGTGATGATGATGATGATGATGATGATGATGAT 1866

PS Example; Fig 9A; 167pp; English.

XX The invention describes a method of targeting a compound to a cell or
CC tissue of an individual expressing a functionally active hyaluronan (HA)
CC receptor for endocytosis (HARE) or a cell that does not express
CC functionally active HARE. The method involves using HA molecule, a
CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
CC a monoclonal antibody raised against a HA-binding domain of HARE. The
CC method is useful for targeting a compound, preferably a chemotherapeutic
CC agent or a radioisotope to cell of an individual, especially a human,
CC expressing HARE on its surface (e.g. gene therapy). Also described is a
CC method useful for preventing interaction between a cell having at least
CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
CC its surface. This second method is useful for preventing metastasis by
CC preventing interaction between tumour cells having HA, CD or CDS coat and
CC non-tumour cells expressing HARE on its surface. The invention also
CC describes a method useful for detecting the presence of HA, CD and CDS in
CC a biological fluid. This sequence encodes the human 190kDa Hyaluronan
CC receptor for endocytosis (HARE)

XX
SQ Sequence 4642 BP; 1141 A; 1215 C; 1216 G; 1070 T; 0 U; 0 Other;

Query Match 99.2%; Score 3232.6; DB 7; Length 4642;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 3238; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY	7	ACAGGAGCTTCCCAAGAACCCGAAACCTCCAGTATTTCTCCAGTTGCAGGAGCATT	56
DB	1394	ATCAGGAGCTTCCCAAGAACCCGAAACCTCCAGTATTTCTTCCAGTTGCAGGAGCATT	1453
QY	67	TGCTGAAGAGATCTGTCGCGCCAGGCCCTTCACTGTTTTTGACCTTTATCTGCAGCCT	126
DB	1454	TGCTGAAGAGATCTGTCGCGCCAGGCCCTTCACTGTTTTTGACCTTTATCTGCAGCCT	1513
QY	127	TTGATGAGGAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGGTTCTTC	186
DB	1514	TTGATGAGGAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGGTTCTTC	1573
QY	187	GGTACCATGTGTCGCTGCCACCGAGCTGCTTCTGGAACCTGAAATGATCTCAATG	246
DB	1574	GGTACCATGTGTCGCTGCCACCGAGCTGCTTCTGGAACCTGAAATGATCTCAATG	1633
QY	247	CTACTTCCCTCAAGGAGAGCCAAATAGTATCATCTCCGCTCTCTCAGAGCAGGTTATATAA	306
DB	1634	CTACTTCCCTCAAGGAGAGCCAAATAGTATCATCTCCGCTCTCTCAGAGCAGGTTATATAA	1693
QY	307	ATAATAAGGCTAAGATCATATCAGTGATATCATCATAGTAAATGAGATGTTTCATATCA	366
DB	1694	ACAATAAGGCTAAGATCATATCAGTGATATCATCATAGTAAATGAGATGTTTCATATCA	1753
QY	367	TAGACAAATTTGCTATCTCCCAAAATTTGCTTATCATCTCCCAAGACAACTCTGGAAGAA	426
DB	1754	TAGACAAATTTGCTATCTCCCAAAATTTGCTTATCATCTCCCAAGACAACTCTGGAAGAA	1813
QY	427	TTCTGCAAAATCTTACGACTTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCACTTAA	486
DB	1814	TTCTGCAAAATCTTACGACTTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCACTTAA	1873
QY	487	TACAGACTCAGGTTTCTGAGTGTATCATCCGATCCATCCACACCCGATCACTCTCT	546
DB	1874	TACAGACTCAGGTTTCTGAGTGTATCATCCGATCCATCCACACCCGATCACTCTCT	1933
QY	547	TCGTGCCCAACGACCAAGCCCTCATGCCCCTACTCTGCTGAACAAACAGGACTTCTCTGTCA	606
DB	1934	TCGTGCCCAACGACCAAGCCCTCATGCCCCTACTCTGCTGAACAAACAGGACTTCTCTGTCA	1993
QY	607	ACCAAGACAAACAGGACCAAGCTGAAGAGTATTTGAAGTTTCATGTGATACGAGATGCCA	666
DB	1994	ACCAAGACAAACAGGACCAAGCTGAAGAGTATTTGAAGTTTCATGTGATACGAGATGCCA	2053
QY	667	AGGTTTGTGTTGGATCTTCCCAACATCCACTGCTGGAGACCCCTGCAAGGTTTCAGAGC	726
DB	2054	AGGTTTGTGTTGGATCTTCCCAACATCCACTGCTGGAGACCCCTGCAAGGTTTCAGAGC	2113

QY	727	TGAGTGTAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTTCTGAATGCCAAACCT	786
DB	2114	TGAGTGTAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTTCTGAATGCCAAACCT	2173
QY	787	GCAGAAATTTGTGCAGCGGAGCTCTTTTGTGACCTCGGTGTGGCCTTACCGCATTTGATGTC	846
DB	2174	GCAGAAATTTGTGCAGCGGAGCTCTTTTGTGACCTCGGTGTGGCCTTACCGCATTTGATGTC	2233
QY	847	TGCTGATTTATCCCACTTGGGGGGCGCTGTGACACCTTTTACTACTTTTCGATGCTCGG	906
DB	2234	TGCTGATTTATCCCACTTGGGGGGCGCTGTGACACCTTTTACTACTTTTCGATGCTCGG	2293
QY	907	GGGAGTGTGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTGG	966
DB	2294	GGGAGTGTGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTGG	2353
QY	967	TGAAGCAGAAGTGTCTCTACAACTGCCCTTCAAGAGGAACCTGGAAGCTGCCGGGAGC	1026
DB	2354	TGAAGCAGAAGTGTCTCTACAACTGCCCTTCAAGAGGAACCTGGAAGCTGCCGGGAGC	2413
QY	1027	GGTGCAGCTGTGTATACAGATCCCAGGTGCTGCAAGGGCTACTTCCGGCGAGACTGTC	1086
DB	2414	GGTGCAGCTGTGTATACAGATCCCAGGTGCTGCAAGGGCTACTTCCGGCGAGACTGTC	2473
QY	1087	AGGCTCTGCCCTGCAGGACCCAGATGCCCGCTGTAAATAACCGGGGTGTCTGCCCTTGATCAGT	1146
DB	2474	AGGCTCTGCCCTGCAGGACCCAGATGCCCGCTGTAAATAACCGGGGTGTCTGCCCTTGATCAGT	2533
QY	1147	ACTCGGCCACCGGAGGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGATGTT	1206
DB	2534	ACTCGGCCACCGGAGGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGATGTT	2593
QY	1207	GCTGCGCGGGAGATTCGGGCTGATGTGTCCTGTCCTGTCAGACACGAGCAGT	1266
DB	2594	GCTGCGCGGGAGATTCGGGCTGATGTGTCCTGTCCTGTCAGACACGAGCAGT	2653
QY	1267	GCGATGATGGCATCACGGGCTCCGGGCTGCTGTGTAACCGGGGTGGACAGGCCCT	1326
DB	2654	GCGATGATGGCATCACGGGCTCCGGGCTGCTGTGTAACCGGGGTGGACAGGCCCT	2713
QY	1327	CGTGTGACACTCAGGAGTGTGCTGAGTGTGTAACCGGCTCTGTCGTCATGCCA	1386
DB	2714	CGTGTGACACTCAGGAGTGTGCTGAGTGTGTAACCGGCTCTGTCGTCATGCCA	2773
QY	1387	CTGTGAAGAGAACACAGCTGTGAGTGTAACTGATTTATGAAGTGAACGAAATCACAT	1446
DB	2774	CTGTGAAGAGAACACAGCTGTGAGTGTAACTGATTTATGAAGTGAACGAAATCACAT	2833
QY	1447	GCACTGTTGATTTCTGCAAAACAGGACAAACCGGGGCTGTGCAAAAGGTGGCCAGATGCT	1506
DB	2834	GCACTGTTGATTTCTGCAAAACAGGACAAACCGGGGCTGTGCAAAAGGTGGCCAGATGCT	2893
QY	1507	CCAGAAAGGACAGAAAGTCTCTGAGCTGCCAGAGGATACAAAGGGAGGGCACA	1566
DB	2894	CCAGAAAGGACAGAAAGTCTCTGAGCTGCCAGAGGATACAAAGGGAGGGCACA	2953
QY	1567	GCTGCACAGATAGACCCCTGTGACAGCGGCTTAAACGAGGGTGTACAGACACGCCA	1626
DB	2954	GCTGCACAGATAGACCCCTGTGACAGCGGCTTAAACGAGGGTGTACAGACACGCCA	3013
QY	1627	CCTGTGAAGATGACAGCGCCGGGCAACAGTGTGAGTGTAAAGTCACTATGTGGAG	1686
DB	3014	CCTGTGAAGATGACAGCGCCGGGCAACAGTGTGAGTGTAAAGTCACTATGTGGAG	3073
QY	1687	ATGGGCTGAATGTGAGCGGAGCAGCTGCCATTCACCGCTCTTACAGGACAAATGGG	1746
DB	3074	ATGGGCTGAATGTGAGCGGAGCAGCTGCCATTCACCGCTCTTACAGGACAAATGGG	3133
QY	1747	AGTGCCATGCAGACGCAAAATGTGTCGACCTCCACTTCCAGGATACCACTTTGGGGTGT	1806
DB	3134	AGTGCCATGCAGACGCAAAATGTGTCGACCTCCACTTCCAGGATACCACTTTGGGGTGT	3193

QY	1807	TCCATCTACGCTCCCACTGGCCAGATATAAGCTTGA	1866
DB	3194	TCCATCTACGCTCCCACTGGCCAGATATAAGCTTGA	3253
QY	1867	GTGCCAAGAAAGCTGCGACCATGCAACTACACCAAGCTCTCTATGCCAGAGGCCA	1926
DB	3254	GTGCCAAGAAAGCTGCGACCATGCAACTACACCAAGCTCTCTATGCCAGAGGCCA	3313
QY	1927	AGTACCAACTGTGCTCAGCAGGCTGGCTGGAGACCGGCGGGTTCCTATGCCACACAGCCT	1986
DB	3314	AGTACCAACTGTGCTCAGCAGGCTGGCTGGAGACCGGCGGGTTCCTATGCCACACAGCCT	3373
QY	1987	TCGCGCTCCAGAACTGTGGCTCTCGTGTGGTTGGGATAGTGGACTATGAGACCTAGACCCA	2046
DB	3374	TCGCGCTCCAGAACTGTGGCTCTCGTGTGGTTGGGATAGTGGACTATGAGACCTAGACCCA	3433
QY	2047	ACAAGAGTGAATGTGGGATGTCTTCCTGCTATCCGATGAAAGATGTGAATCTGACCTGTATGT	2106
DB	3434	ACAAGAGTGAATGTGGGATGTCTTCCTGCTATCCGATGAAAGATGTGAATCTGACCTGTATGT	3493
QY	2107	AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACTGCTGCAGGTCTCTGATGT	2166
DB	3494	AGGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACTGCTGCAGGTCTCTGATGT	3553
QY	2167	CCTTCCCTCACTCACAAACTTCTGACAGGAAGTGTGTGGGCTATTCGAAAGCTCAGCTC	2226
DB	3554	CCTTCCCTCACTCACAAACTTCTGACAGGAAGTGTGTGGGCTATTCGAAAGCTCAGCTC	3613
QY	2227	GAGGCCGTGCAATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGACACCTCTTTGTGC	2286
DB	3614	GAGGCCGTGCAATTTCTAGAACACCTGACTGACCTGTCCATCCGCGGACACCTCTTTGTGC	3673
QY	2287	CACAGAAACAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGCGGGAATCGAGCACCA	2346
DB	3674	CACAGAAACAGTGGGCTGGGGAGAAATGAGACCTTGTCTGGGCGGGAATCGAGCACCA	3733
QY	2347	TCGCCAATGTGAGCATGTTTCTCAATGACCTTGTCAATGGACCAACCTGCAACCGA	2406
DB	3734	TCGCCAATGTGAGCATGTTTCTCAATGACCTTGTCAATGGACCAACCTGCAACCGA	3793
QY	2407	GGGTGGGAAGCAAGCTGCTCATCTGCCAGCAGGACCCACTTCCACCGACGGAGCCA	2466
DB	3794	GGCTGGGAAGCAAGCTGCTCATCTGCCAGCAGGACCCACTTCCACCGACGGAGCCA	3853
QY	2467	GGTTTGTGTGAGGAAGCAATCTGCAGTGGGACATCTTTTGCTCCAAATGGGATCATTC	2526
DB	3854	GGTTTGTGTGAGGAAGCAATCTGCAGTGGGACATCTTTTGCTCCAAATGGGATCATTC	3913
QY	2527	ATGTCATTTCCAGGCCCTTTAAAGCACCCCTGCCCCGCTGACCTTGACCCACATGGCT	2586
DB	3914	ATGTCATTTCCAGGCCCTTTAAAGCACCCCTGCCCCGCTGACCTTGACCCACATGGCT	3973
QY	2587	TGGGAGCAGGAGTCTCTTTGGCATCATCTGCTGTGACTGGGGCTGTGCTTGGCTGCTT	2646
DB	3974	TGGGAGCAGGAGTCTCTTTGGCATCATCTGCTGTGACTGGGGCTGTGCTTGGCTGCTT	4033
QY	2647	ACTCCTACTTTCGGATAAACCAGGAGAACAAATCGGCTTCCAGACATTTTGTAGTCGGAAGAG	2706
DB	4034	ACTCCTACTTTCGGATAAACCAGGAGAACAAATCGGCTTCCAGACATTTTGTAGTCGGAAGAG	4093
QY	2707	ACATTAAATGTGAGCTCTTTGGCAAGCAGAGCCTGAGAAATATCTCGAACCCCTTGTATG	2766
DB	4094	ACATTAAATGTGAGCTCTTTGGCAAGCAGAGCCTGAGAAATATCTCGAACCCCTTGTATG	4153
QY	2767	AGAGCAAACTCAGCTCCCCAGAACCTTCTACGACCCCTTCAAGGACTCTGGAAGAAC	2826
DB	4154	AGAGCAAACTCAGCTCCCCAGAACCTTCTACGACCCCTTCAAGGACTCTGGAAGAAC	4213
QY	2827	GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCTGACCGGAGATGCCAG	2886
DB	4214	GGCAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGGCTGACCGGAGATGCCAG	4273
QY	2887	CCATCACTCACTGCCACTGGGCCATCAACTGTGAATTTCTCAGCACCAAGTTGCTTTTATG	2946

Db	4274	CCATCACTCACTGCCACCTGGGCCATCAACTGTGAATTCTCAGCACCAAGTTGCCTTTTAG	4333
Qy	2947	GAACGTAAGTCCTTTTAAGCACTCAGAAAGCCATCACTCTCTCGCTGATCTGGGG	3006
Db	4334	GAACGTAAGTCCTTTTAAGCACTCAGAAAGCCATCACTCTCTCGCTGATCTGGGG	4393
Qy	3007	TTGTTTCTGTGGGTGAGAGATGCTTTGCTGTGCCACCCAGTACAGCTTCCTCTCTCTGAC	3066
Db	4394	TTGTTTCTGTGGGTGAGAGATGCTTTGCTGTGCCACCCAGTACAGCTTCCTCTCTCTGAC	4453
Qy	3067	CCTTTGGCTCTTCTTCCCTTTTGTACTTTCAGCTGGCACCTGCTCCATCTTCGCCCTACATG	3126
Db	4454	CCTTTGGCTCTTCTTCCCTTTTGTACTTTCAGCTGGCACCTGCTCCATCTTCGCCCTACATG	4513
Qy	3127	ATGGGTAACTGTGATCACTTCTTCCCTGTTAGATTGTAAGCCTCCTCTTTGTTATCCACG	3186
Db	4514	ATGGGTAACTGTGATCACTTCTTCCCTGTTAGATTGTAAGCCTCCTCTTTGTTATCCACG	4573
Qy	3187	CCCTAGCCCACTGCTGCACACAGAACTGTGCACAATAAAGGTTTATGGAAACAGAAAAA	3246
Db	4574	CCCTAGCCCACTGCTGCACACAGAACTGTGCACAATAAAGGTTTATGGAAACAGAAAAA	4633
Qy	3247	AAAAAAA 3254	
Db	4634	AGTCAACA 4641	
RESULT 4			
ABT08487			
XX	ID	ABT08487 standard; cDNA; 8444 BP.	
AC	AC	ABT08487;	
XX	XX		
DT	28-NOV-2002	(first entry)	
XX	XX		
DE	Human novel protein NOV1a coding sequence SEQ ID NO: 1.		
XX	XX		
KW	Human; NOVX;	single nucleotide polymorphism; SNP; anti-HIV; cytostatic;	
KW	antiarteriosclerotic;	antidiabetic; antiasthmatic; antiinflammatory;	
KW	haemostatic; hypotensive;	neuroprotective; anorectic; nootropic;	
KW	antidepressant;	immunosuppressive; antibacterial; antiparasitic;	
KW	virucide; tranquilizer;	anticonvulsant; osteopathic; analgesic;	
KW	antiparkinsonian;	dermatological; antiinfertility; cerebroprotective;	
KW	antiaddictive; gene; ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FH	Key	Location/Qualifiers	
FT	variation	replace(4185,C)	
FT		/*tag= a	
FT	variation	replace(4524,C)	
FT		/*tag= b	
FT	variation	replace(4654,A)	
FT		/*tag= c	
FT	variation	replace(4671,G)	
FT		/*tag= d	
FT	variation	replace(4820,C)	
FT		/*tag= e	
XX	XX		
PN	WO200259315-A2.		
XX	XX		
PD	01-AUG-2002.		
XX	XX		
PF	19-DEC-2001;	2001WO-US050076.	
XX	XX		
PR	19-DEC-2000;	2000US-0256619P.	
PR	19-JAN-2001;	2001US-0262959P.	
PR	28-FEB-2001;	2001US-0272408P.	
PR	20-APR-2001;	2001US-0285189P.	
PR	26-JUL-2001;	2001US-0308039P.	
PR	09-AUG-2001;	2001US-0311266P.	
XX	XX		

(CURA-) CURAGEN CORP.

PA Shinkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;
PI Shenoy S, Sphytek KA, Gangolli E, Miller C, Boldog F, Li L;
PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
PI Rothenberg M;
XX WPI; 2002-666903/71.
DR P-PSDB; ABU10586.
XX New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders e.g.
PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
PT or Alzheimer's disease.

Claim 8; Page 17-19; 363pp; English.

XX The present invention provides the protein and coding sequences of
CC several novel human proteins, designated NOVX. These can be used in the
CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
CC Lindau syndrome, Alzheimer's disease, stroke, tuberous sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesh-Nyhan syndrome, multiple sclerosis, ataxia
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
CC graft-versus-host disease. The present sequence is a coding sequence of
CC the invention

XX Sequence 8444 BP; 2251 A; 2105 C; 2147 G; 1929 T; 0 U; 12 Other;

Query Match 92.0%; Score 2999.8; DB 6; Length 8444;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 3128; Conservative 0; Mismatches 18; Indels 108; Gaps 1;

QY 7 ACAGAGCTTCCCAAGACCCGAAACCTCCAGTATTCTCCAGTTGCCAGGACATT 66
DB 5279 ATCAGAGCTTCCCAAGACCCGAAACCTCCAGTATTCTCCAGTTGCCAGGACATT 5338
QY 67 TCGTGAAGATCTGGTCCGCCAGGCCCTTCACTGTTTTCACCTTTATCTGCAGCCT 126
DB 5339 TCGTGAAGATCTGGTCCGCCAGGCCCTTCACTGTTTTCACCTTTATCTGCAGCCT 5398
QY 127 TTGATCAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCAGGTTCTTC 186
DB 5399 TTGATCAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCAGGTTCTTC 5458
QY 187 GGTACCATGTGGTCCGCTGCCACAGCTGTTCTGGAAACCTGAAATTGATCTCAATG 246
DB 5459 GGTACCATGTGGTCCGCTGCCACAGCTGTTCTGGAAACCTGAAATTGATCTCAATG 5518
QY 247 CTACTTCCCTCCCAAGAGACCAATAGTCACTCCGCTCTCTCAGAGCAGGTTATATAA 306
DB 5519 CTACTTCCCTCCCAAGAGACCAATAGTCACTCCGCTCTCTCAGAGCAGGTTATATAA 5578
QY 307 ATAAAGGCTAAGATCATATCAGTGATATCATAGTACTAATGGATTGTTTATATCA 366
DB 5579 ATAAAGGCTAAGATCATATCAGTGATATCATAGTACTAATGGATTGTTTATATCA 5638
QY 367 TAGACAAATGGCTATCTCCCAAAATTTGCTTATCATCTCCAAAGACAACTCTGGAGAA 426
DB 5639 TAGACAAATGGCTATCTCCCAAAATTTGCTTATCATCTCCAAAGACAACTCTGGAGAA 5698
QY 427 TTCTGCAAAATCTTACGACTTTGGCAACAAATGGCTACATCAAAATTTAGCAACTTAA 486
DB 5699 TTCTGCAAAATCTTACGACTTTGGCAACAAATGGCTACATCAAAATTTAGCAACTTAA 5758
QY 487 TACAGGACTCAGGTTTGGTGTGATCATACCGATCCCATCCACCCAGTCACTCTCT 546
DB 5759 TACAGGACTCAGGTTTGGTGTGATCATACCGATCCCATCCACCCAGTCACTCTCT 5818

QY 547 TCTGCCCCACGACCAAGCCCTCCATGCTACCTGCTGAAACAAAGAGACTTCTCTGTCA 606
DB 5819 TCTGCCCCACGACCAAGCCCTCCATGCTACCTGCTGAAACAAAGAGACTTCTCTGTCA 5878
QY 607 ACCAAGACAAACGACCAAGCTGAAAGAGTATTTCAAGTTCATGTATACAGATGCCA 666
DB 5879 ACCAAGACAAACGACCAAGCTGAAAGAGTATTTCAAGTTCATGTATACAGATGCCA 5938
QY 667 AGGTTTATAGTGTGATCTTCCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 726
DB 5939 AGGTTTATAGTGTGATCTTCCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 5998
QY 727 TCAGTGTGAAATGTGGAGCTGCGAGGACATCCGTCACCTCTTTCTGAATGCCAAACCT 786
DB 5999 TCAGTGTGAAATGTGGAGCTGCGAGGACATCCGTCACCTCTTTCTGAATGCCAAACCT 6058
QY 787 GCAGAAATGTGCGAGGGAGCTCTTTGTGACCTGGGTGGCCCTACCGCATGATGTGC 846
DB 6059 GCAGAAATGTGCGAGGGAGCTCTTTGTGACCTGGGTGGCCCTACCGCATGATGTGC 6118
QY 847 TCGTGTATGATCCCAACCTGGGGGGCGCTGTGACACCTTTACTTTCGATGCTCCG 906
DB 6119 TCGTGTATGATCCCAACCTGGGGGGCGCTGTGACACCTTTACTTTCGATGCTCCG 6178
QY 907 GGGAGTGTGGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG 966
DB 6179 GGGAGTGTGGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG 6238
QY 967 TGAAGCAGAAAGTGTCTTCAACCTGCCCTTCAAGAGGAACTTGAAGAGCTGCCGGAGC 1026
DB 6239 TGAAGCAGAAAGTGTCTTCAACCTGCCCTTCAAGAGGAACTTGAAGAGCTGCCGGAGC 6298
QY 1027 GGTGCGACCTGTGTATACAGATCCCGAGTGTGCAAGGGCTACTTCGCGCGAGACTGTC 1086
DB 6299 GGTGCGACCTGTGTATACAGATCCCGAGTGTGCAAGGGCTACTTCGCGCGAGACTGTC 6358
QY 1087 AGGCTTCCCTCGAGGACCAAGATGCCCTCGTAAATAACCGGGGTGTCTGCTTTGATCAGT 1146
DB 6359 AGGCTTCCCTCGAGGACCAAGATGCCCTCGTAAATAACCGGGGTGTCTGCTTTGATCAGT 6418
QY 1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGCTTCAATGGGACGGGTGTGAGATGT 1206
DB 6419 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGCTTCAATGGGACGGGTGTGAGATGT 6478
QY 1207 GCTGCGCGGGAGATTCGGGCTGATTGCTGCTGCTGCTGCTCAGACACCGACAGT 1266
DB 6479 GCTGCGCGGGAGATTTGGGCTGATTGCTGCTGCTGCTGCTCAGACACCGACAGT 6538
QY 1267 GCGATGATGGCATCAGGGCTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1326
DB 6539 GCGATGATGGCATCAGGGCTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6598
QY 1327 CGTGTGACACTCAGGAGATTTGCTGCGAGTGTGACCGCTCTCTTCTGCTCATGCCA 1386
DB 6599 CGTGTGACACTCAGGAGATTTGCTGCGAGTGTGACCGCTCTCTTCTGCTCATGCCA 6658
QY 1387 CCTGTAAGAGAACACACGTTGAGTGTAACTCGATTATGAAGGTGACGGAATCACAT 1446
DB 6659 CCTGTAAGAGAACACACGTTGAGTGTAACTCGATTATGAAGGTGACGGAATCACAT 6718
QY 1447 GCACAGTTTGTGGATTTCTGCAACAGGACAAACCGGGGCTGTGCAAGGTGGCCAGATGCT 1506
DB 6719 GCACAGTTTGTGGATTTCTGCAACAGGACAAACCGGGGCTGTGCAAGGTGGCCAGATGCT 6778
QY 1507 CCAGAGAGGACAGAGTCTCTGCGAGTGTGCAAGGGATACAAAGGGGACGGGCACA 1566
DB 6779 CCAGAGAGGACAGAGTCTCTGCGAGTGTGCAAGGGATACAAAGGGGACGGGCACA 6838
QY 1567 GCTGCAAGAGATAGACCTCTGTGCAAGCGGCTTAAACGGAGGGTGTACAGACACGCCA 1626
DB 6839 GCTGCAAGAGATAGACCTCTGTGCAAGCGGCTTAAACGGAGGGTGTACAGACACGCCA 6898

QY 1627 CCTTAAAGATGACAGCCCGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGGAG 1686
Db 6899 CTTGTAAAGATGACAGCCCGGCAAGCAAGTGTGAGTGTAAAGTCACTATGTCGGAG 6958
QY 1687 ATGGCTGAACCTGTGAGCGGAGAGCTGCCCAATTGACCGCTGCTTACAGGACAATGGGC 1746
Db 6959 ATGGCTGAACCTGTGAGCGGAGAGCTGCCCAATTGACCGCTGCTTACAGGACAATGGGC 7018
QY 1747 AGTGCCATGACAGCCCAAAATGTGCGACCTCACTTCCAGGATACCACTGTGGGGTGT 1806
Db 7019 AGTGCCATGACAGCCCAAAATGTGCGACCTCACTTCCAGGATACCACTGTGGGGTGT 7078
QY 1807 TCCATCTAGCTCCCACTGGCCCAAGTATAAGCTGACCTTTGACAAAGCCAGAGGGCT 1866
Db 7079 TCCATCTAGCTCCCACTGGCCCAAGTATAAGCTGACCTTTGACAAAGCCAGAGGGCT 7138
QY 1867 GTGCCAAGAAAGCTCGACCAATGGCAACCTACAACAGCTCTCTTATGCCAGAAAGGCCA 1926
Db 7139 GTGCCAAGAAAGCTCGACCAATGGCAACCTACAACAGCTCTCTTATGCCAGAAAGGCCA 7198
QY 1927 AGTACCACTGTGCTCAGAGGCTGGCTGGAGACCGGGGGGTTGCCCTACCCCAAGGCT 1986
Db 7199 AGTACCACTGTGCTCAGAGGCTGGCTGGAGACCGGGGGGTTGCCCTACCCCAAGGCT 7258
QY 1987 TCGCTCCCAAGAACTGTGGGATGCTTCTGCTATCGGATGAAGATGTGAATGCACTGCA 2046
Db 7259 TCGCTCCCAAGAACTGTGGGATGCTTCTGCTATCGGATGAAGATGTGAATGCACTGCA 7318
QY 2047 ACAAGAGTGAATGTGGGATGCTTCTGCTATCGGATGAAGATGTGAATGCACTGCA 2106
Db 7319 ACAAGAGTGAATGTGGGATGCTTCTGCTATCGGATGAAGATGTGAATGCACTGCA 7358
QY 2107 AGGTGGGCTATGTGGGAGATGCTTCTCATGCAGTGGGAACTGTGCTGAGGTCCTGTGT 2166
Db 7359 ----- 7358
QY 2167 CTTTCCCTCACTACAAACTTCTGTGACGAGTGTGCTGCTATTCACACAGCTCAGCTC 2226
Db 7359 -----GGAAGTGTGGCTTATTCACACAGCTCAGCTC 7390
QY 2227 GAGCGGTGCAATTTCTAGAACAACCTGACTGCTCCATCCGCGGACCCCTTTTGTGC 2286
Db 7391 GAGCGGTGCAATTTCTAGAACAACCTGACTGCTCCATCCGCGGACCCCTTTTGTGC 7450
QY 2287 CACAGAACAGTGGCTGGGGGAGATGAGACCTTGTCTGGGGGAGACATCGAGACCAACC 2346
Db 7451 CACAGAACAGTGGCTGGGGGAGATGAGACCTTGTCTGGGGGAGACATCGAGACCAACC 7510
QY 2347 TCGCAATGTGAGCATGTTTTTCTACAATGACCTTGTCAATGGGACACCCCTGCAACGA 2406
Db 7511 TCGCAATGTGAGCATGTTTTTCTACAATGACCTTGTCAATGGGACACCCCTGCAACGA 7570
QY 2407 GGGTGGGAAGCAAGTGTCTATCACTGCGCAGCAGGACCCACTCCAAACCGCAGGAGACA 2466
Db 7571 GGGTGGGAAGCAAGTGTCTATCACTGCGCAGCAGGACCCACTCCAAACCGGTAACAGTA 7630
QY 2467 GGTGTTGTGATGGAAGACCAATTCGAGTGGGACATCTTTGCCCTCCAAATGGGATCATTC 2526
Db 7631 GGTGTTGTGATGGAAGACCAATTCGAGTGGGACATCTTTGCCCTCCAAATGGGATCATTC 7690
QY 2527 ATGTCAATTTCCAGGCTTTAAAGACACCCCTGCCCCGCTGACCTTGACCCACACTGGCT 2586
Db 7691 ATGTCAATTTCCAGGCTTTAAAGACACCCCTGCCCCGCTGACCTTGACCCACACTGGCT 7750
QY 2587 TGGGAGCAGGATCTTCTTTGCCATCATCTGTGTGACTGGGGCTGTTGCCCTTGGCTGCTT 2646
Db 7751 TGGGAGCAGGATCTTCTTTGCCATCATCTGTGTGACTGGGGCTGTTGCCCTTGGCTGCTT 7810
QY 2647 ACTCTACTTTCCGATTAACCGGAGAACAAATCGGCTTCCAGCAATTTTGTGTCGGAAGAG 2706
Db 7811 ACTCTACTTTCCGATTAACCGGAGAACAAATCGGCTTCCAGCAATTTTGTGTCGGAAGAG 7870
QY 2707 ACATTAATGTTGACGCTCTTGGCAAGCAGCAGCTGAGAATATCTCGNAACCCCTTGTATG 2766

Db 7871 ACATTAATGTTGACGCTCTTGGCAAGCAGCAGCTGAGAATATCTCGNAACCCCTTGTATG 7930
QY 2767 AGAGCACAACTCAGCTCCCGCAGAACCTTCTCAGACCCCTTACGGACTCTGAAGAAC 2826
Db 7931 AGAGCACAACTCAGCTCCCGCAGAACCTTCTCAGACCCCTTACGGACTCTGAAGAAC 7990
QY 2827 GGCAGCTTGGAGGCAATGACCCCTTGGAGCACATCTGAGGGCTGGACGGAGATGCCAG 2886
Db 7991 GGCAGCTTGGAGGCAATGACCCCTTGGAGCACATCTGAGGGCTGGACGGAGATGCCAG 8050
QY 2887 CCATCACTCAGTCCCACTGGGCCATCAACTGTGTAATCTCAGACCAAGTTCCTTTAG 2946
Db 8051 CCATCACTCAGTCCCACTGGGCCATCAACTGTGTAATCTCAGACCAAGTTCCTTTAG 8110
QY 2947 GAACTAAAGTCTTTAAAGCACTCAGAGCCATCACTCTCTGCTGCTGATCTGGGCTGGGG 3006
Db 8111 GAACTAAAGTCTTTAAAGCACTCAGAGCCATCACTCTCTGCTGCTGATCTGGGCTGGGG 8170
QY 3007 TTGTTTCTGCTGGTGAGAGATGTTGCTGTGCCCAACCCAGTACAGCTTCTCTCTGAC 3066
Db 8171 TTGTTTCTGCTGGTGAGAGATGTTGCTGTGCCCAACCCAGTACAGCTTCTCTCTGAC 8230
QY 3067 CCTTTGGCTCTTCTCTCTTTGCTTCTCAGTGGCACCTGCTCAATCTGCTGCTTACATG 3126
Db 8231 CCTTTGGCTCTTCTCTCTTTGCTTCTCAGTGGCACCTGCTCAATCTGCTGCTTACATG 8290
QY 3127 ATGGGTAAGTGTGATTTTCTTCTTCTGCTGTTAGATGTTAGACCTTCTTGTATCCCCAG 3186
Db 8291 ATGGGTAAGTGTGATTTTCTTCTTCTGCTGTTAGATGTTAGACCTTCTTGTATCCCCAG 8350
QY 3187 CCCTAGCCAGTGCCTGACACAGGAACCTGTGCACAATAAGGTTTATGGAAACAGAAAAA 3246
Db 8351 CCCTAGCCAGTGCCTGACACAGGAACCTGTGCACAATAAGGTTTATGGAAACAGAAAAA 8410
QY 3247 AAAAAAAGAAAAA 3260
Db 8411 AGTCAAAAAAAGAAAA 8424
RESULT 5
ABT08488
ID ABT08488 standard; cDNA; 8495 BP.
XX AC ABT08488;
XX AC
XX DT 28-NOV-2002 (first entry)
XX DE Human novel protein NOV1b coding sequence SEQ ID NO: 3.
XX DE Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
KW antiarteriosclerotic; antidiabetic; antilasthmatic; antiinflammatory;
KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
XX antiaddictive; gene; ss.
OS Homo sapiens.
XX WO200259315-A2.
XX 01-AUG-2002.
XX 19-DEC-2001; 2001WO-US050076.
XX 19-DEC-2000; 2000US-0256619P.
PR 19-JAN-2001; 2001US-0262959P.
PR 28-FEB-2001; 2001US-0272408P.
PR 20-APR-2001; 2001US-0285189P.
PR 26-JUL-2001; 2001US-0308039P.
PR 09-AUG-2001; 2001US-0311266P.
XX

PA	(CURA-) CURAGEN CORP.	
XX	Shinkets RA, Patturajan M, Vernet CM, Casman SJ, Malyankar U;	
PI	Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;	
PI	Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;	
PI	Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;	
PI	Rothenberg M;	
XX	WPI; 2002-666903/71.	
DR	P-PSDB; ABU10587.	
XX		
XX	New isolated NOVX polypeptides and polynucleotides, useful for	
PT	preventing, diagnosing or treating NOVX-associated disorders e.g.	
PT	diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease	
PT	or Alzheimer's disease.	
XX		
PS	Claim 8; Page 21-23; 363pp; English.	
XX		
CC	The present invention provides the protein and coding sequences of	
CC	several novel human proteins, designated NOVX. These can be used in the	
CC	treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-	
CC	Lindau syndrome, Alzheimer's disease, stroke, tuberous sclerosis,	
CC	hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral	
CC	palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia	
CC	telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,	
CC	obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,	
CC	infertility, atherosclerosis, hypertension, scleroderma, haemophilia,	
CC	diabetes, pancreatitis, autoimmune disease, asthma, arthritis,	
CC	immunodeficiencies, HIV, viral, bacterial or parasitic infections, or	
CC	graft-versus-host disease. The present sequence is a coding sequence of	
CC	the invention	
XX		
SQ	Sequence 8495 BP; 2295 A; 2102 C; 2135 G; 1957 T; 0 U; 6 Other;	
	Query Match 91.9%; Score 2995.6; DB 6; Length 8495;	
	Best Local Similarity 96.1%; Pred. No. 0;	
	Matches 3137; Conservative 0; Mismatches 10; Indels 116; Gaps 2;	
QY	7 AGCAGAGCTTCCAGAGCCGAAACTCCAGTATTTCTCCAGTTGCGAGGACATT 66	
DB	5311 ATCAGAGCTTCCAGAGCCGAAACTCCAGTATTTCTCCAGTTGCGAGGACATT 5370	
QY	67 TCGTGAAGATCTGGTCGCGCCAGGCCCTTCACTGTTTGTGACCTTTATCTGACGCT 126	
DB	5371 TCGTGAAGATCTGGTCGCGCCAGGCCCTTCACTGTTTGTGACCTTTATCTGACGCT 5430	
QY	127 TTGATCAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGGTTCTTC 186	
DB	5431 TTGATCAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGGTTCTTC 5490	
QY	187 GGTACCATGTGTCGCTGCCAGCAGTGTCTTGGAAACCTGAAATGATCTCAAATG 246	
DB	5491 GGTACCATGTGTCGCTGCCAGCAGTGTCTTGGAAACCTGAAATGATCTCAAATG 5550	
QY	247 CTACTTCCCTCCAGGAGCCCAATAGTATCTCCGCTCTCTCAGACGAGGTTATATAA 306	
DB	5551 CTACTTCCCTCCAGGAGCCCAATAGTATCTCCGCTCTCAGACGAGGTTATATAA 5610	
QY	307 ATAATAAGGCTAAGATCATATCCAGTATATCATATCATATAGTAAATGGATTTGTTTATCA 366	
DB	5611 ACAATAAGGCTAAGATCATATCCAGTATATCATATCATATAGTAAATGGATTTGTTTATCA 5670	
QY	367 TAGACAAATGCTATCTCCAAAATTTGCTTATCATCTCCAAAGACACTCTGGAGAA 426	
DB	5671 TAGACAAATGCTATCTCCAAAATTTGCTTATCATCTCCAAAGACACTCTGGAGAA 5730	
QY	427 TTCTGCAAAATCTTACGACTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCAACTTA 486	
DB	5731 TTCTGCAAAATCTTACGACTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCAACTTA 5790	
QY	487 TACAGACTCAGGTTTCTGAGTGTATCATCCGATCCATCCACCCAGTCACTCTCT 546	
DB	5791 TACAGACTCAGGTTTCTGAGTGTATCATCCGATCCATCCACCCAGTCACTCTCT 5850	
QY	547 TCTGCCCAACCGACCAA-----GCCCTCATGCTCCTACTCTGTGAACAACAGGACT 597	
DB	5851 TCTGCCCAACCGACCAAAGCCCTCATGCTCCTCATGCTGCTGACACAACAGGACT 5910	
QY	598 TCTCTTTCAACAACAAGACAACAAGCAAGCTGAAGAGTATTGAAAGTTTCATGTGATAC 657	
DB	5911 TCTCTTTCAACAACAAGACAACAAGCAAGCTGAAGAGTATTGAAAGTTTCATGTGATAC 5970	
QY	658 GAGATGCCAAGGTTTACTGTGGATCTTCCACATCCACTGCTGCTGGAAGACCTTCACAG 717	
DB	5971 GAGATGCCAAGGTTTACTGTGGATCTTCCACATCCACTGCTGCTGGAAGACCTTCACAG 6030	
QY	718 GTTCAGAGCTGAGTGTGAAATGTGAGCTGGCAGGACATCCGTCACCTTCCTTCTGAAATG 777	
DB	6031 GTTCAGAGCTGAGTGTGAAATGTGAGCTGGCAGGACATCCGTCACCTTCCTTCTGAAATG 6090	
QY	778 GCCAAACCTGCAGAAATGTGCAGCGGAGCTCTTGTGTGACCTGGGTGGGCTACGGCA 837	
DB	6091 GCCAAACCTGCAGAAATGTGCAGCGGAGCTCTTGTGTGACCTGGGTGGGCTACGGCA 6150	
QY	838 TTGACTGTCTGCTGATGATCCACCTGGGGGGCGCTGTGACACCTTTACTACTTTCG 897	
DB	6151 TTGACTGTCTGCTGATGATCCACCTGGGGGGCGCTGTGACACCTTTACTACTTTCG 6210	
QY	898 ATGCCTCGGGGAGTGTGGGAGCTGTGTCAATACTCCAGCTGCCCAAGGTGGAGTAAAC 957	
DB	6211 ATGCCTCGGGGAGTGTGGGAGCTGTGTCAATACTCCAGCTGCCCAAGGTGGAGTAAAC 6270	
QY	958 CAAAGGTTGTGAAGCAGAGTGTCTTCAACCTGCGCCTTGAAGAGAACCTGGGAAGGCT 1017	
DB	6271 CAAAGGTTGTGAAGCAGAGTGTCTTCAACCTGCGCCTTGAAGAGAACCTGGGAAGGCT 6330	
QY	1018 GCGGGAGCGGTGCAGCCTGGTGATACAGATCCCGAGGTGCTGCAAGGGCTACTTCGGGC 1077	
DB	6331 GCGGGAGCGGTGCAGCCTGGTGATACAGATCCCGAGGTGCTGCAAGGGCTACTTCGGGC 6390	
QY	1078 GAGACTGTGAGCCTGCGCTGGAGGACCAAGATGCCCCGCTGTAAATACCGGGGTGTCTGCC 1137	
DB	6391 GAGACTGTGAGCCTGCGCTGGAGGACCAAGATGCCCCGCTGTAAATACCGGGGTGTCTGCC 6450	
QY	1138 TTGATCAGTACTCGGCCACCGGAGAGTGAATGAAACACCGGCTTCAATGGGACGGCGT 1197	
DB	6451 TTGATCAGTACTCGGCCACCGGAGAGTGAATGAAACACCGGCTTCAATGGGACGGCGT 6510	
QY	1198 GTGAGATGTGTCGCGCGGGAGATTGCGGCTGATTTGCTGCTGCTGCTGCTGCTGACAC 1257	
DB	6511 GTGAGATGTGTCGCGCGGGAGATTGCGGCTGATTTGCTGCTGCTGCTGCTGCTGACAC 6570	
QY	1258 ACGGACAGTGCATGATGTCATCACCGGCTCCGGGAGTGCCTCTCTGAAACCGGGGTGGA 1317	
DB	6571 ACGGACAGTGCATGATGTCATCACCGGCTCCGGGAGTGCCTCTCTGAAACCGGGGTGGA 6630	
QY	1318 CAGGCCCCCTGCTGTCAGTCAAGGAGTTTGGCTCAGTGTGTACGCTCTCTGTTCTG 1377	
DB	6631 CAGGCCCCCTGCTGTCAGTCAAGGAGTTTGGCTCAGTGTGTACGCTCTCTGTTCTG 6690	
QY	1378 CTCATGCCACCTGTAGAGGAGAACACACGTTGAGTGAACCTGGATTTATGAAGTGAAC 1437	
DB	6691 CTCATGCCACCTGTAGAGGAGAACACACGTTGAGTGAACCTGGATTTATGAAGTGAAC 6750	
QY	1438 GAATCATGTCACAGTGTGTGGATTTCTGCAAAACAGGACAAACCGGGGCTGTGCAAGGTTG 1497	
DB	6751 GAATCATGTCACAGTGTGTGGATTTCTGCAAAACAGGACAAACCGGGGCTGTGCAAGGTTG 6810	
QY	1498 CAGATGTCTCCAGAGAGGACAGAGTCTCTGAGCTGCCAGAGGAGTATCAAAAGGG 1557	
DB	6811 CAGATGTCTCCAGAGAGGACAGAGTCTCTGAGCTGCCAGAGGAGTATCAAAAGGG 6870	
QY	1558 ACGGGCAGAGTGCACAGAGATAGACCCCTGTGTCAGACCGGCTTAAACGGAGGTTGTCAAC 1617	
DB	6871 ACGGGCAGAGTGCACAGAGATAGACCCCTGTGTCAGACCGGCTTAAACGGAGGTTGTCAAC 6930	

QY	1618	AGCAGCCACCTGTAAAGATGACAGCCCGGCAAGCAACAAGTGTAGTGTAAAAAGTCACT	1671
Db	6931	AGCAGCCACCTGTAAAGATGACAGCCCGGCAAGCAACAAGTGTAGTGTAAAAAGTCACT	6990
QY	1678	ATGTCCGAGATGGGCTGAACCTGTGAGCCGGAGACGCTGCCCAATTGACCGCTGCTTACAGG	1737
Db	6991	ATGTCCGAGATGGGCTGAACCTGTGAGCCGGAGACGCTGCCCAATTGACCGCTGCTTACAGG	7050
QY	1738	ACAATGGGCACTGGCATGCAGACGCGMAATGTGTGCACCTCCACTTCCAGGATACCACTG	1797
Db	7051	ACAATGGGCACTGGCATGCAGACGCGMAATGTGTGCACCTCCACTTCCAGGATACCACTG	7110
QY	1798	TTGGGGTGTTCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCA	1857
Db	7111	TTGGGGTGTTCATCTACGCTCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCCA	7170
QY	1858	GAGAGGCTGTGCCAAAGAGCTGGACCATGGCAACTACAAACCAAGCTCTCCTATGCCC	1917
Db	7171	GAGAGGCTGTGCCAAAGAGCTGGACCATGGCAACTACAAACCAAGCTCTCCTATGCCC	7230
QY	1918	AGAAGGCCAAGTACCACCTGTGTCTCAGCAGGCTGGGACCGGCGGGTGTGCCCTACC	1977
Db	7231	AGAAGGCCAAGTACCACCTGTGTCTCAGCAGGCTGGGACCGGCGGGTGTGCCCTACC	7290
QY	1978	CCACAGCCTTCGCCCTCCCAAGAACTGTGGCTCTGGTGTGGTGGGATAGTGGACTATGGAC	2037
Db	7291	CCACAGCCTTCGCCCTCCCAAGAACTGTGGCTCTGGTGTGGTGGGATAGTGGACTATGGAC	7350
QY	2038	CTACACCCCAACAGAGTGAAATGTGGGATGCTCTCTGCTATCGGATCAAGATGTGAACT	2097
Db	7351	CTACACCCCAACAGAGTGAAATGTGGGATGCTCTCTGCTATCGGATCAAGATGTGAACT	7400
QY	2098	GCACCTGCAGAGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACTGCTGCAGG	2157
Db	7401	GCACCTGCAGAGTGGGCTATGTGGGAGATGGCTTCTCATGCAGTGGGAACTGCTGCAGG	7400
QY	2158	TCCTGTATGTCCTTCCCTCCTACTCAAACTTCCTGCAGGAAGTGTGGCCCTATTCCAACA	2217
Db	7401	TCCTGTATGTCCTTCCCTCCTACTCAAACTTCCTGCAGGAAGTGTGGCCCTATTCCAACA	7423
QY	2218	GCTCAGCTCGAGGCCGTGCATTTCTAGAAACACTGACTGACCTGTGCATCGGCGGACCC	2277
Db	7424	GCTCAGCTCGAGGCCGTGCATTTCTAGAAACACTGACTGACCTGTGCATCGGCGGACCC	7483
QY	2278	TCCTTGTGCCACAGAAACAGTGGGCTGGGGAGAAATGAGACTTGTCTCGGCGCGGACATCG	2337
Db	7484	TCCTTGTGCCACAGAAACAGTGGGCTGGGGAGAAATGAGACTTGTCTCGGCGCGGACATCG	7543
QY	2338	AGCACCACTCGCCAAATGTGCAGATGTTTTTTCTACAATGACCTTGTCAATGGCACCAACC	2397
Db	7544	AGCACCACTCGCCAAATGTGCAGATGTTTTTTCTACAATGACCTTGTCAATGGCACCAACC	7603
QY	2398	TGCAAAACGAGGCTGGGAAGCAAGCTGCTCATCACTGCAGCAGGACCCACTCCAAACCGA	2457
Db	7604	TGCAAAACGAGGCTGGGAAGCAAGCTGCTCATCACTGCAGCAGGACCCACTCCAAACCGA	7663
QY	2458	CGGAGACCAAGTGTGTGATGGGAAGGCCATTCTGCAGTGGGACATCTTTGGCTCCAAATG	2517
Db	7664	CGGAGACCAAGTGTGTGATGGGAAGGCCATTCTGCAGTGGGACATCTTTGGCTCCAAATG	7723
QY	2518	GGATCATTCATGCTACTTTCCAGGCTTTAAAGACACCCCTGCCCGCGTGACCTTGACCC	2577
Db	7724	GGATCATTCATGCTACTTTCCAGGCTTTAAAGACACCCCTGCCCGCGTGACCTTGACCC	7783
QY	2578	ACACTGGCTTGGGAGCAGGATCTTCTTTGGCATCATCTGGTGTGACTGGGGCTGTGGCT	2637
Db	7784	ACACTGGCTTGGGAGCAGGATCTTCTTTGGCATCATCTGGTGTGACTGGGGCTGTGGCT	7843
QY	2638	TGGCTGCTTACTCTACTTTCCGATAAACCGGAGAACAAATCGGCTTCCAGCATTTTGAGT	2697
Db	7844	TGGCTGCTTACTCTACTTTCCGATAAACCGGAGAACAAATCGGCTTCCAGCATTTTGAGT	7903
QY	2698	CGGAAGAGGACATTAATGTGTTGCAGCTCTTGGCAAGCAGCAGCCTGAGAATATCTCGAAC	2757

Db	7904	CGAAGAGGACATTAAATGTTGCAGCTCTTGGCAAGCAGCGCTTGAGATATCTCGAAC	7963
Qy	2758	CTTGTATGAGAGCACAACCTCAGCTCCCCAGAACTTCTTCTACGACCCCTTTCACGAGCT	2817
Db	7964	CTTGTATGAGAGCACAACCTCAGCTCCCCAGAACTTCTTCTACGACCCCTTTCACGAGCT	8023
Qy	2818	CTGAAGAACGCGAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGCGCTTGGACGG	2877
Db	8024	CTGAAGAACGCGAGCTTGAGGGCAATGACCCCTTGAGGACACTGTGAGGCGCTTGGACGG	8083
Qy	2878	AGATGCCAGCCATCACTCAGCTGCCACCTTGCGCCATCAACTGTGAATTTCTCAGCACCACTT	2937
Db	8084	AGATGCCAGCCATCACTCAGCTGCCACCTTGCGCCATCAACTGTGAATTTCTCAGCACCACTT	8143
Qy	2938	GCCTTTTAGGAACGTAAGTCTTTAAGCACTCAGAGCCATCACTCATCTCTCTGGCTG	2997
Db	8144	GCCTTTTAGGAACGTAAGTCTTTAAGCACTCAGAGCCATCACTCATCTCTCTGGCTG	8203
Qy	2998	ATCTGGGGGTGTTTCTGTGGGTGAGAGATGTGTGCTGTGCCACCCAGTACAGCTTCC	3057
Db	8204	ATCTGGGGGTGTTTCTGTGGGTGAGAGATGTGTGCTGTGCCACCCAGTACAGCTTCC	8263
Qy	3058	TCCTCTGACCCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACTGCTCCATTCTG	3117
Db	8264	TCCTCTGACCCCTTTGGCTCTTCTTCCTTTGTACTCTTCAGCTGGCACTGCTCCATTCTG	8323
Qy	3118	CCCTACATGATGGTAACTGTGATCTTCTCCCTGTGTAGATTGTAAGCTCCNCTTTG	3177
Db	8324	CCCTACATGATGGTAACTGTGATCTTCTCCCTGTGTAGATTGTAAGCTCCNCTTTG	8383
Qy	3178	TATCCAGCCCTAGCCAGTGCCTGACAGAGAACTGTGCACAAATAAAGTTTATGGAA	3237
Db	8384	TATCCAGCCCTAGCCAGTGCCTGACAGAGAACTGTGCACAAATAAAGTTTATGGAA	8443
Qy	3238	CAGAAAAAATAAAAAAAAAAAAAA 3260	
Db	8444	CAGAAAAAATAAAAAAAAAAAAAA 8466	
RESULT 6			
ABX13823			
ID	ABX13823 standard; cDNA; 4962 BP.		
AC	ABX13823;		
DT	19-FEB-2003 (first entry)		
DE	cDNA encoding human 190kDa Hyaluronan receptor for endocytosis #2.		
KW	Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;		
KW	chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;		
XX	tumour; gene therapy; human; gene; ss.		
OS	Homo sapiens.		
PH	Key		
FT	CDS		
FT	Location/Qualifiers		
FT	1..4962		
FT	/*tag= a		
FT	/product= "HARE"		
FT	/note= "Hyaluronan receptor for endocytosis"		
FT	/partial		
FT	/note= "No start codon given"		
XX			
FN	WO200286093-A2.		
XX			
PD	31-OCT-2002.		
XX			
PF	25-APR-2002; 2002WO-US013209.		
XX			
PR	25-APR-2001; 2001US-00842930.		
XX			
PR	25-APR-2001; 2001US-0286468P.		
XX			

PA (WEIG/) WEIGEL P H.
PA (WEIG/) WEIGEL J A.
XX Weigel PH, Weigel JA;
XX P-PSDB; ABG72514.
XX
PT Targeting compounds e.g. chemotherapeutic agent to cell of subject
PT expressing functional active hyaluronan receptor for endocytosis of HARE,
PT by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
PT epitope.
XX
XX Example; Fig 9B; 167pp; English.
XX
CC The invention describes a method of targeting a compound to a cell or
CC tissue of an individual expressing a functionally active hyaluronan (HA)
CC receptor for endocytosis (HARE) or a cell that does not express
CC functionally active HARE. The method involves using HA molecule, a
CC chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
CC a monoclonal antibody raised against a HA-binding domain of HARE. The
CC method is useful for targeting a compound, preferably a chemotherapeutic
CC agent or a radioisotope to cell of an individual, especially a human,
CC expressing HARE on its surface (e.g. gene therapy). Also described is a
CC method useful for preventing interaction between a cell having at least
CC one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
CC its surface. This second method is useful for preventing metastasis by
CC preventing interaction between tumour cells having HA, CD or CDS coat and
CC non-tumour cells expressing HARE on its surface. The invention also
CC describes a method useful for detecting the presence of HA, CD and CDS in
CC a biological fluid. This sequence encodes a longer version of the human
CC 190kDa Hyaluronan receptor for endocytosis (HARE) shown in ABG72499
CC (Encoded by AEX13822)
XX
SQ Sequence 4962 BP; 1249 A; 1277 C; 1311 G; 1125 T; 0 U; 0 Other;

Query Match 87.5%; Score 2853.2; DB 7; Length 4962;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2855; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 AGCAGAGCTTCCCAAGAACCCGAAACATTCACAGTATTTCTTCCAGTTTGCAGGAGCATTT 66
DB 2105 ATCAGAGCTTCCCAAGAACCCGAAACATTCACAGTATTTCTTCCAGTTTGCAGGAGCATTT 2164

QY 67 TCGTGAAGATCTGTCGCGCCAGCCCTTCTACCTGTTTTCACCTTTTATCTGACGCT 126
DB 2165 TCGTGAAGATCTGTCGCGCCAGCCCTTCTACCTGTTTTCACCTTTTATCTGACGCT 2224

QY 127 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCGAGGTTCTTC 186
DB 2225 TTGATGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCGAGGTTCTTC 2284

QY 187 GGTACCATGTGTCGCTCGCAGCTGCTTCTGGAACACCTGAATTTGATCTCAATG 246
DB 2285 GSTACCATGTGTCGCTCGCAGCTGCTTCTGGAACACCTGAATTTGATCTCAATG 2344

QY 247 CTACTTCCCTCCAAAGAGACCAATAGTATCATCTCCGCTCTCAGACACGGTGTATATAA 306
DB 2345 CTACTTCCCTCCAAAGAGACCAATAGTATCATCTCCGCTCTCAGACACGGTGTATATAA 2404

QY 307 ATAATAAGGCTAAGATCATATCATATCATATCATATCATATCATATCATATCATATCA 366
DB 2405 ACRATAAGGCTAAGATCATATCATATCATATCATATCATATCATATCATATCATATCA 2464

QY 367 TAGACAAATGCTATCTCCGAAATTTGCTTATCATCTCCGAAAGACCACTCTGGAAGAA 426
DB 2465 TAGACAAATGCTATCTCCGAAATTTGCTTATCATCTCCGAAAGACCACTCTGGAAGAA 2524

QY 427 TTCTGCAAAATCTTAGACTTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCACTTAA 486
DB 2525 TTCTGCAAAATCTTAGACTTTTGGCAACAAACAAATGGCTACATCAAAATTTAGCACTTAA 2584

QY 487 TACAGGACTCAGGTTTGTGCTGAGTGTATCATCCGATCCCATCCACACCCAGTCACTCTCT 546

DB 2585 TACAGGACTCAGGTTTGTGCTGAGTGTATCATCCGATCCCATCCACACCCAGTCACTCTCT 2644

QY 547 TCTGGCCCCACGACCAAGCCCTCCATGCGCTTACCTGCTGAACAACAGGACTTCTCTGTTC 606

DB 2645 TCTGGCCCCACGACCAAGCCCTCCATGCGCTTACCTGCTGAACAACAGGACTTCTCTGTTC 2704

QY 607 ACCAAGACCAACAGGACCAAGCTGAAGAGTATTTTGAAGTTTTCATGTATACAGATGCCA 666

DB 2705 ACCAAGACCAACAGGACCAAGCTGAAGAGTATTTTGAAGTTTTCATGTATACAGATGCCA 2764

QY 667 AGGTTTACGCTGTGGATCTTCCACATCCACTGCTGGAAGACCTTGAAGTTTCAGAGC 726

DB 2765 AGGTTTACGCTGTGGATCTTCCACATCCACTGCTGGAAGACCTTGAAGTTTCAGAGC 2824

QY 727 TGAGTGTAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTGTAATGCCAAACCT 786

DB 2825 TGAGTGTAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTGTAATGCCAAACCT 2884

QY 787 GCAGAAATGTGACGCGGAGCTCTTGTGTTTGAACCTGGGTGTGGCCTACGSCATTTGACTGTC 846

DB 2885 GCAGAAATGTGACGCGGAGCTCTTGTGTTTGAACCTGGGTGTGGCCTACGSCATTTGACTGTC 2944

QY 847 TGCTGATTTATCCCAACCTTGGGGGCGCTGTGACACCTTTACTTACTTTCGATGCTCGG 906

DB 2945 TGCTGATTTATCCCAACCTTGGGGGCGCTGTGACACCTTTACTTACTTTCGATGCTCGG 3004

QY 907 GGGAGTGTGGAGCTGTCTCAATATCTCCAGCTGCCAAGGTGGAGTAAACCAAGGGTG 966

DB 3005 GGGAGTGTGGAGCTGTCTCAATATCTCCAGCTGCCAAGGTGGAGTAAACCAAGGGTG 3064

QY 967 TGAAGCAGAAGTGTCTTACAACTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGAGC 1026

DB 3065 TGAAGCAGAAGTGTCTTACAACTGCCCTTCAAGAGGAACCTGGAAGGCTGCCGGAGC 3124

QY 1027 GGTGACGCTGTGTATACAGATCCCAAGTGTCTCAAGGGCTTACTTTGGGCGAGACTGTC 1086

DB 3125 GGTGACGCTGTGTATACAGATCCCAAGTGTCTCAAGGGCTTACTTTGGGCGAGACTGTC 3184

QY 1087 AGGCTTGCCTCGAGGACCAAGTCCCGGTAAATAACCGGGGTGTCTCCCTTGATCACT 1146

DB 3185 AGGCTTGCCTCGAGGACCAAGTCCCGGTAAATAACCGGGGTGTCTCCCTTGATCACT 3244

QY 1147 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGGCTGTGAGATGT 1206

DB 3245 ACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGGCTGTGAGATGT 3304

QY 1207 GGTGCGGGGAGATTCGGGCTGATTTGTCTGCCCTGTGGCTGCTCAGACCGGACAGT 1266

DB 3305 GGTGCGGGGAGATTCGGGCTGATTTGTCTGCCCTGTGGCTGCTCAGACCGGACAGT 3364

QY 1267 GCGATGATGGCATCAGGGCTCGGGCAGTGTCTGTGAAACGGGGTGGACAGGCCCT 1326

DB 3365 GCGATGATGGCATCAGGGCTCGGGCAGTGTCTGTGAAACGGGGTGGACAGGCCCT 3424

QY 1327 CGTGTGACACTCAGGAGTTCCTCGCTGAGTGTGACGCTCTTGTTCGTCTGATGCCA 1386

DB 3425 CGTGTGACACTCAGGAGTTCCTCGCTGAGTGTGACGCTCTTGTTCGTCTGATGCCA 3484

QY 1387 CTTGTAGGAGAACCAACAGTGTGATGTAACTGTGATTTAGAGTGTGAGGATTCACAT 1446

DB 3485 CTTGTAGGAGAACCAACAGTGTGATGTAACTGTGATTTAGAGTGTGAGGATTCACAT 3544

QY 1447 GCACAGTTCGTGATTTCTCAAAACAGGACAAACGGGGCTGTGCAAAAGTGGCCAGATGCT 1506

DB 3545 GCACAGTTCGTGATTTCTCAAAACAGGACAAACGGGGCTGTGCAAAAGTGGCCAGATGCT 3604

QY 1507 CCCAGAGGGCACGAGAGTCTCTGTGAGTGTGAGAGGATCAAAAGGGGACGGGACCA 1566

DB 3605 CCCAGAGGGCACGAGAGTCTCTGTGAGTGTGAGAGGATCAAAAGGGGACGGGACCA 3664

QY 1567 GCTGACACAGATAGACCTTGTGACAGGGCTTAAACGGAGGGTGTCTGAGGACGACCA 1626

CC patients own production of protein. They are used to produce NOVX
CC proteins, by inserting the nucleic acid into a cell and culturing it to
CC express the protein. The DNA may be used as DNA probes in assays to
CC detect and quantitate the presence of similar DNAs in samples, and which
CC patients may need restorative therapy. The NOVX protein may also be used
CC as antigens in the production of antibodies (Abs) against NOVX and in
CC assays to identify modulators of NOVX expression and activity. The anti-
CC NOVX Abs and antagonist are used to down regulate expression and
CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX
CC in samples. Disorders that may be prevented, diagnosed and/or treated
CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
CC NOV13-16 have casein kinase II phosphorylation sites characteristic of
CC serine/threonine kinases, and are used to treat kinase-related disorders
CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
CC -like super family and are involved in, e.g. regulation of cell
CC development, apoptosis, cell adhesion, growth migration, cell structure
CC and motility and protein management, and are used to treat cancers,
CC inflammatory disorders, immune disorders and cellular adhesion disorders.
CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
CC syndrome and marfan syndrome
XX
SQ Sequence 3625 BP; 896 A; 971 C; 914 G; 844 T; 0 U; 0 Other;

Query Match 86.8%; Score 2828.8; DB 4; Length 3625;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2985; Conservative 0; Mismatches 43; Indels 111; Gaps 2;
QY 122 AGCCTTTGATGAGGAAGCTCGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGGT 181
DB 569 ACCCTCTGATCTCTGTACAGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCCAGGT 628
QY 182 TCTTCGGTACCAATGGTTCGCTGCGCCGACCAAGCTGCTTCTGGAAAACTGAAATTTGATCTC 241
DB 629 TCTTCGGTACCAATGGTTCGCTGCGCCGACCAAGCTGCTTCTGGAAAACTGAAATTTGATCTC 688
QY 242 AAATGCTACTTCCCTCAAGGAGAGCCAAATAGTATCTCCGCTCTCAGAGCAGCGTGTA 301
DB 689 AAATGCTACTTCCCTCAAGGAGAGCCAAATAGTATCTCCGCTCTCAGAGCAGCGTGTA 748
QY 302 TATAAATAAAGGCTAAGATCATATCCAGTGTATATCATCAGTACTAATGGGATTTGTCA 361
DB 749 TATAAATAAAGGCTAAGATCATATCCAGTGTATATCATCAGTACTAATGGGATTTGTCA 808
QY 362 TATCATAGACAAATGCTATCTCCCAAAATTTGCTTTATCATCTCCCAAGACAACTCTGG 421
DB 809 TATCATAGACAAATGCTATCTCCCAAAATTTGCTTTATCATCTCCCAAGACAACTCTGG 868
QY 422 AAGAAATCTGCAAAATCTTACGACTTTGGCAACAACAATGGCTACATCAAAATTTAGCAA 481
DB 869 AAGAAATCTGCAAAATCTTACGACTTTGGCAACAACAATGGCTACATCAAAATTTAGCAA 928
QY 482 CTTTAATACAGGACTCAGGTTTGTGAGTGTATCATCCGATCCCATCCACACCCCAAGTCAC 541
DB 929 CTTTAATACAGGACTCAGGTTTGTGAGTGTATCATCCGATCCCATCCACACCCCAAGTCAC 988
QY 542 TCTCTTCTGCCCCACGACCAAGCCCTCCATGCGCTTACCTGTGAAACAAGCAAGCTTCCT 601
DB 989 TCTCTTCTGCCCCACGACCAAGCCCTCCATGCGCTTACCTGTGAAACAAGCAAGCTTCCT 1048
QY 602 GTTCAACCAAGACAAAGGAGAGCTGAAGGAGTATTTGAAAGTTTTCATGTGATACGAGA 661
DB 1049 GTTCAACCAAGACAAAGGAGAGCTGAAGGAGTATTTGAAAGTTTTCATGTGATACGAGA 1108
QY 662 TGCCAAAGGTTTATGCTGTGATCTTCCCAATCCATGCTGGAAGACCCCTGCAAGGTTTC 721
DB 1109 TGCCAAAGGTTTATGCTGTGATCTTCCCAATCCATGCTGGAAGACCCCTGCAAGGTTTC 1168
QY 722 AGAGCTGAGTGTGAAATGTGAGCTGGCAGGAGCATCGGTGACCTCTTTCTGATGACCA 781
DB 1169 AGAGCTGAGTGTGAAATGTGAGCTGGCAGGAGCATCGGTGACCTCTTTCTGATGACCA 1228

QY 782 AACCTGCAGAAATGTGCAGCGGAGAGCTCTTGTGTTTACCTGGTGTGGCTACGGCAATGA 841
DB 1229 AACCTGCAGAAATGTGCAGCGGAGAGCTCTTGTGTTTACCTGGTGTGGCTACGGCAATGA 1288
QY 842 CTGCTCTGATGATGATCCACCTTGGGGGGCGGTGTGACACCTTTACTCTTTTCGATGC 901
DB 1289 CTGCTCTGATGATGATCCACCTTGGGGGGCGGTGTGACACCTTTACTCTTTTCGATGC 1348
QY 902 CTCGGGGAGGTGGGAGCTGTGTCAATCTCCAGCTGCCCAAGGTGGAGTAAACCCAAA 961
DB 1349 CTCGGGGAGGTGGGAGCTGTGTCAATCTCCAGCTGCCCAAGGTGGAGTAAACCCAAA 1408
QY 962 GGGTGTGAAGCAGAAAGTGTCTCTACAACCTGCTTCAAGAGGAACTTGAAGGCTGCGG 1021
DB 1409 GGGTGTGAAGCAGAAAGTGTCTCTACAACCTGCTTCAAGAGGAACTTGAAGGCTGCGG 1468
QY 1022 GGAGCGGTGCAGCCTGTGTGATACAGATCCCAGGTGTCTGCAAGGGCTACTTCGGGCGAGA 1081
DB 1469 GGAGCGGTGCAGCCTGTGTGATACAGATCCCAGGTGTCTGCAAGGGCTACTTCGGGCGAGA 1528
QY 1082 CTGTGAGGCTGCTGCTGAGGAGACAGATGCCCTGTGTAATACCGGGGTGTCTGCTTGA 1141
DB 1529 CTGTGAGGCTGCTGCTGAGGAGACAGATGCCCTGTGTAATACCGGGGTGTCTGCTTGA 1584
QY 1142 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACCGGCTGTA 1201
DB 1585 TCAGTACTCGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACCGGCTGTA 1644
QY 1202 GATGTGCTGGCGGGGAGATTCGGGCTGATGTGCTGCCCTGTGGCTGTCTCAGACCAACG 1261
DB 1645 GATGTGCTGGCGGGGAGATTCGGGCTGATGTGCTGCCCTGTGGCTGTCTCAGACCAACG 1704
QY 1262 ACAGTGCAGATGATGGCATCACGGGCTCCGGGCGAGTGCCTCTGTGAAACGGGGTGGACAG 1321
DB 1705 ACAGTGCAGATGATGGCATCACGGGCTCCGGGCGAGTGCCTCTGTGAAACGGGGTGGACAG 1764
QY 1322 CCCCTGCTGTGACACTCAGCGAGTGTTCCTGCGAGTGTAGCCTCTCTTGTCTCTCTCA 1381
DB 1765 CCCCTGCTGTGACACTCAGCGAGTGTTCCTGCGAGTGTAGCCTCTCTTGTCTCTCTCA 1824
QY 1382 TGGCACTGTAAAGGAGAAACAACGCTGTGAGTGTAACTGGATTAAGAGGTGACCGAAT 1441
DB 1825 TGGCACTGTAAAGGAGAAACAACGCTGTGAGTGTAACTGGATTAAGAGGTGACCGAAT 1884
QY 1442 CACATGCACAGTTGTGGATTTCTGCAAAACAGGACAAACGGGGCTGTGCAAAAGGTGGCCAG 1501
DB 1885 CACATGCACAGTTGTGGATTTCTGCAAAACAGGACAAACGGGGCTGTGCAAAAGGTGGCCAG 1944
QY 1502 ATGCTCCCAAGAGGGCAACGAGTCTCTGCGAGTCCAGAGAGGGATACAAAGGGACGG 1561
DB 1945 ATGCTCCCAAGAGGGCAACGAGTCTCTGCGAGTCCAGAGAGGGATACAAAGGGACGG 2004
QY 1562 GCACAGCTGCACAGATAGACCCCTGTGCGAGCGGCTTAAACGAGGGTGTCCACAGCA 1621
DB 2005 GCACAGCTGCACAGATAGACCCCTGTGCGAGCGGCTTAAACGAGGGTGTCCACAGCA 2064
QY 1622 CGCCACCTGTAAAGATCACAGGCGCGGGCAAGCAAAAGTGTGAGTGTAAAGTCACTATGT 1681
DB 2065 CGCCACCTGTAAAGATCACAGGCGCGGGCAAGCAAAAGTGTGAGTGTAAAGTCACTATGT 2124
QY 1682 CGGAGATGGGCTGAACCTGTGAGCGGAGCAGCTGCCCATTTGACCGCTGTACAGGACAA 1741
DB 2125 CGGAGATGGGCTGAACCTGTGAGCGGAGCAGCTGCCCATTTGACCGCTGTACAGGACAA 2184
QY 1742 TGGGAGTGCATGCGAGACGCCAAATGTGTGCACTCCACTTCCAGGATACCACTGTGG 1801
DB 2185 TGGGAGTGCATGCGAGACGCCAAATGTGTGCACTCCACTTCCAGGATACCACTGTGG 2244
QY 1802 GGTGTTCCATCTACGCTGCCACTTGGGCGAGTAAAGTGAACCTTTGACAAAGCCAGAGA 1861
DB 2245 GGTGTTCCATCTACGCTGCCACTTGGGCGAGTAAAGTGAACCTTTGACAAAGCCAGAGA 2304
QY 1862 GGCTGTGCCAACGAGAGCTGCGACCATGGCAACCTCAACACAGCTCTCTCTATGCCAGAA 1921

```
|||||
Db 2305 GGCCTGTGCAACGAAGCTGGACCATGCAACCTACACCAAGCTCTCTATGCCAGAA 2364
QY 1922 GGCACAGTACCACTGTCTGACAGGCTGGTGGAGACCGGGCGGTTGCCCTACCCAC 1981
Db 2365 GGCACAGTACCACTGTCTGACAGGCTGGTGGAGACCGGGCGGTTGCCCTACCCAC 2424
QY 1982 AGCCTTCGCTCCCAAACTGTGGCTCTGGTGGTGGAGTAGTGACATATGGACCTAG 2041
Db 2425 AGCCTTCGCTCCCAAACTGTGGCTCTGGTGGTGGAGTAGTGACATATGGACCTAG 2484
QY 2042 ACCCAACAAGAGTGAATGTGGATGTCTTCTGCTATCGGATGAAAGATGTGAATGCAC 2101
Db 2485 ACCCAACAAGAGTGAATGTGGATGTCTTCTGCTATCGGATGAA----- 2530
QY 2102 CTGCAAGTGGGCTATGTGGAGATGGCTTCTCATGCAAGTGGAAACCTGCTGCAAGTCT 2161
Db 2531 ----- 2530
QY 2162 GATGTCTTCCCTCACTCACAACTTCTGAGGAGTGTCTGCTATTCCAAAGCTC 2221
Db 2531 -----GGAAGTGTGGCTATTCCAAAGCTC 2557
QY 2222 AGCTCAGGCGCTGATTTCTAGAACACCTGACTGACCTGTCCATCCGGCGCACCTCTT 2281
Db 2558 AGCTCAGGCGCTGATTTCTAGAACACCTGACTGACCTGTCCATCCGGCGCACCTCTT 2617
QY 2282 TGTGCCACAGAACAGTGGGCTGGGGAGATGAGACCTTGTCTGGCGGGACATCGAGCA 2341
Db 2618 TGTGCCACAGAACAGTGGGCTGGGGAGATGAGACCTTGTCTGGCGGGACATCGAGCA 2677
QY 2342 CCACCTCCCAATGTGACATGTTTTCTACAAATGACCTTGTCAATGACACACCTTGA 2401
Db 2678 CCACCTCCCAATGTGACATGTTTTCTACAAATGACCTTGTCAATGACACACCTTGA 2737
QY 2402 AACGAGGCTGGGAAGCAAGCTGCTCATCACTGCCAGCAGGACCACTCCAAACCGACGA 2461
Db 2738 AACGAGGCTGGGAAGCAAGCTGCTCATCACTGCCAGCAGGACCACTCCAAACCGACGA 2797
QY 2462 GACCAAGTTTGTGTAGTAAGAGCAATCTTGTGAGGAGCAATCTTGTGCTCCAAATGGAT 2521
Db 2798 GACCAAGTTTGTGTAGTAAGAGCAATCTTGTGAGGAGCAATCTTGTGCTCCAAATGGAT 2857
QY 2522 CATTCATGCTCATTTCCAGGCTTTAAAGACCCCTGCGCCGCTGACCTTGAACCCAC 2581
Db 2858 CATTCATGCTCATTTCCAGGCTTTAAAGACCCCTGCGCCGCTGACCTTGAACCCAC 2917
QY 2582 TGGCTTGGGACAGGATCTTCTTTGCCATCATCTGTGTGACTGGGGCTGTGCTTGGC 2641
Db 2918 TGGCTTGGGACAGGATCTTCTTTGCCATCATCTGTGTGACTGGGGCTGTGCTTGGC 2977
QY 2642 TGCTTACTCTTACTTGGATTAACCGAGAACATCGGCTTCCAGCAATTTGAGTGGGA 2701
Db 2978 TGCTTACTCTTACTTGGATTAACCGAGAACATCGGCTTCCAGCAATTTGAGTGGGA 3037
QY 2702 AGAGGACATTAATGTGTGAGCTCTTGGCAAGCAGCAGCTGAGATATCTCGAACCCCTT 2761
Db 3038 AGAGGACATTAATGTGTGAGCTCTTGGCAAGCAGCAGCTGAGATATCTCGAACCCCTT 3097
QY 2762 GTATGAGAGCAACCTCAGCTCCCGCAGAACCTTCTTACGACCCCTTACCGACTCTGA 2821
Db 3098 GTATGAGAGCAACCTCAGCTCCCGCAGAACCTTCTTACGACCCCTTACCGACTCTGA 3157
QY 2822 AGAAGGAGCTTGGGGCAATGACCTTGGAGACACTGTGAGGCGCTGGACGGAGAT 2881
Db 3158 AGAAGGAGCTTGGGGCAATGACCTTGGAGACACTGTGAGGCGCTGGACGGAGAT 3217
QY 2882 GCCAGGCACTCACTCACTGCCACCTGGGCACTCACTGTGAATTTCTGAGCACCAGTTGCCT 2941
Db 3218 GCCAGGCACTCACTCACTGCCACCTGGGCACTCACTGTGAATTTCTGAGCACCAGTTGCCT 3277
QY 2942 TTTAGGAACGTAAAGTCTTTTAAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCT
|||||
```

```
3278 TTTAGGAACGTAAAGTCTTTTAAAGCACTCAGAAGCCATACCTCATCTCTCTGGCTGATCT 3337
3002 GGGGGTGTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCCACCCAGTACAGCTTCTCTCT 3061
3338 GGGGGTGTGTTTCTGTGGGTGAGAGATGTGTTGCTGTGCCCCACCCAGTACAGCTTCTCTCT 3397
3062 CTGACCCCTTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3121
3398 CTGACCCCTTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3457
3122 ACATCATGCGGTAACTGTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3181
3458 ACATCATGCGGTAACTGTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3517
3182 CCAGGCCCTTAGCCAGTGTGCTGACACAGGAACCTGTGCAAGAACTGTGCAATAAAGGTTTATGGAACAGA 3241
3518 CCAGGCCCTTAGCCAGTGTGCTGACACAGGAACCTGTGCAAGAACTGTGCAATAAAGGTTTATGGAACAGA 3577
3242 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3260
3578 AACAAAGTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3596

RESULT 8
ID ABA04648 standard; cDNA; 4706 BP.
XX
AC ABA04648;
XX
DT 22-FEB-2002 (first entry)
XX
DE Rat Hyaluronic Acid Receptor for Endocytosis, HARE, coding sequence.
XX
DE HARE; Hyaluronic Acid Receptor for Endocytosis; hyaluronan; chondroitin;
KW chondroitin sulphate; extracellular matrix; cartilage; skin;
KW vitreous humour; endocytic receptor; glycosaminoglycan; rat; ss.
XX
OS Rattus norvegicus.
XX
Key Location/Qualifiers
CDS 1..4296
FT /*tag= a
FT /partial
FT /product= "Rat HARE"
FT /note= "No start codon given"
PN WO200181544-A2.
PD
PP 01-NOV-2001.
XX
PP 25-APR-2001; 2001WO-US013403.
XX
PP 25-APR-2000; 2000US-0199538P.
PR 02-NOV-2000; 2000US-0245320P.
XX
XX (WEIG// WEIGEL P A.
PA (ZHOU// ZHOU B.
PA (WEIG// WEIGEL J A.
XX
XX Weigel PA, Zhou B, Weigel JA;
XX
XX WPI; 2002-049271/06.
DR P-PSDB; AAM47675.
XX
XX New mammalian hyaluronic acid receptor for endocytosis, useful e.g. for
PT identifying agents that inhibit binding to hyaluronic acid, and related
PT nucleic acid.
XX
XX Claim 49; Fig 21; 263pp; English.
PS
CC The present invention relates to sequences for rat and human HARE
CC (Hyaluronic Acid Receptor for Endocytosis, ABA04648, ABA04662, AAM47675
CC and AAM47684). HARE can bind specifically to at least one of hyaluronic
```


CC acid (HA, also known as hyaluronan), chondroitin (C) or chondroitin
CC sulphate (CS). HA is an extracellular matrix component of all tissues, in
CC particular cartilage, skin and vitreous humor. HARE is the endocytic
CC receptor responsible for removing HA and other glycosaminoglycans from
CC the circulation. The present sequence is the coding sequence for rat HARE
XX
SQ Sequence 4706 BP; 1169 A; 1273 C; 1230 G; 1034 T; 0 U; 0 Other;

Query Match 59.4%; Score 1936.8; DB 6; Length 4706;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 652; Indels 14; Gaps 6;

QY 11 GGAGCTTCCAGAACCGGAAACCTCCAGTATTTCTCCAGTTGCGAGGACATTTGCT 70
DB 1395 GGAGCTTCCAGAACCGGAAACCTCCAGTATTTCTCCAGTTGCGAGGACATTTGCT 1454
QY 71 GAAAGATCTGGTCGCGCCAGGCGCCCTTCACTGTTTTTGACCTTTATCTGCAGCCCTTGA 130
DB 1455 CCGAGAGCTTGCTGACCTGGCCCTTCACTGTTGGCCCTTGTCTAGCTCCTTCAA 1514
QY 131 TGAGGAAGCTCGGGTTAAAGACTGGGACAAATACGGTTTAAATGCCCGCAGGTTCTTCGGTA 190
DB 1515 TCATGAGCCCGGATTAAGACTGGGATCAGCAGGCGCTCATGTCCAGGTTCTTCGCTA 1574
QY 191 CGATGTGTCGCTGCCACAGCTGTTCTGGAACCTGAAATGATCTCAATGCTAC 250
DB 1575 TCACGTGTGGCTGCGCAGCAGCTGCTGTGGACAACTTAAAGTGACACAGTGCCAC 1634
QY 251 TTCCCTCCAGGAGGACCAATAGTCACTTCCGTCTCTCAGACACGGTGTATATAATAA 310
DB 1635 GACCTCCAGGAGGACGATTTCCATCTCTCTCAGGACACTGTGTTCAATAACAA 1694
QY 311 TAAGGCTAAGATCATATCCAGTGATATCATAGTACTATAGGATGTTTCATATCATAGA 370
DB 1695 TGAGGCGAAGTCTGTCCAGTGACATCATCAGCAACCAATGGCGTCACTCCAGTTATAGA 1754
QY 371 CAAATGTCTATCTCCAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGATTTCT 430
DB 1755 CAAATGTCTCTCCAAATTTGCTTATCACTCCCAAGACAACTCTGGAAGGTTCT 1814
QY 431 GCAAAATCTTAGACTTTGGCAACAAATAGTCTACATCAAAATTTAGCAACTTAATACA 490
DB 1815 GCAAAATCTTAGACTTTGGCAACAAATAGTCTACATCAAAATTTAGCAACTTAATACA 1874
QY 491 GGAATCAGGTTTGTGAGTGATATCAGGATCCATCCAGTCCATCCAGTCACTCTCTCTG 550
DB 1875 GGAATCAGGTTTGTGAGTGATATCAGGATCCATCCAGTCCATCCAGTCACTCTCTCTG 1934
QY 551 GCCACCGACCAAGCCCTCATGCCCTACCTGCTGAAACACAGCACTTCTGTTCAACCA 610
DB 1935 GCTTACGGAACAGCCCTGGAAGCCTTGCCCCAGAGCAGCAGGACTTCTGTTCAATCA 1994
QY 611 AGACAAACAGGACAAGCTGAAGGATTTGAAGTTTCAATGATACGAGATGCCAAGGT 670
DB 1995 AGACAAACAGGACAAGCTGAAGGATTTGAAGTTTCAATGATACGAGATGCCAAGGT 2054
QY 671 TTTAGCTGTGATTTCTCCACATCCAGTCTGCTGGAGACCTCGAAGTTTCAGAGCTGAG 730
DB 2055 TTTAGCTGTGATTTCTCCACATCCAGTCTGCTGGAGACCTCGAAGTTTCAGAGCTGAG 2114
QY 731 TGTGAATGTGAGCTGGCAGGACATCGGTGACCTCTTCTGAATGGCCAAACCTGCAG 790
DB 2115 TGTGAATGTGAGCTGGCAGGACATCGGTGACCTCTTCTGAATGGCCAAACCTGCAG 2174
QY 791 AATTGTGACGCGGAGCTTTGTTTGAACCTGGGTGTGGCTACGGCAATGACTGTGCT 850
DB 2175 ATTCAATACCGGGAATCTTTGTTTGAACCTGGGTGTGGCTACGGCAATGACTGTGCT 2234
QY 851 GATTGATCCACCTCGGGGCGCTGTGACACCTTTACTACTTTCCGATGCTCGGGGA 910
DB 2235 CATGAATCCACCTCGGGGCGCTGTGACACCTTTACTACTTTCCGATGCTCGGGGA 2294
QY 911 GTGTGGGAGCTGTGTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGGTGAA 970

DB 2295 GTGCGAAGTTGCAATTTTCACTCCAAATGCCCACTGAAGAGCAAGCCAAAGGGCGTGAA 2354
QY 971 CGAAGAGTGTCTCTACAACC---TCCCCCTTCAAGAGGAACCTTGAAGAGCTGCCGGAGCG 1027
DB 2355 GAAGAAGTGTATCTACAAACCCGTTTACCTTTACCTTTACCTTTACCTTTACCTTTACCTTT 2414
QY 1028 GTGCAAGCTGTGTATACAGATCCCGAGGTGTGCAAGGGCTACTTTCCGGCGGAGACTGTCA 1087
DB 2415 GTGCAAGCTGTGTATACAGATCCCGAGGTGTGCAAGGGCTACTTTACCTTTACCTTTACCTTT 2474
QY 1088 GGCCTGCTCGAGGACAGATGCCCTCGTGAATTAACCGGGGTGTCTGCTTGTATCAGTA 1147
DB 2475 GGCCTGCTCGAGGACAGATGCCCTCGTGAATTAACCGGGGTGTCTGCTTGTATCAGTA 2534
QY 1148 CTCGGCCACCGGAGAGTGTAAATGCAACACCGCTTCAATGGGACGGGTGTGAGATGTG 1207
DB 2535 CACACCCATGGGACAGTGCCTATGCCACACCGCTTCAACCGGACAGCTGCGAGCTCTG 2594
QY 1208 CTGGCCGGGAGATTCGGGCTCTGATTTGTCTGCTCTGTGGCTGCTCAGACCAACCGGACAGTG 1267
DB 2595 CTGGCATGGGAGATTTGGGCTCTGACTGTGACCCCGCAGCTGCTCCGAGCATGACAGTG 2654
QY 1268 CGATGATGSCATCAAGGGCTCCGGGCTGCTGCTCTGTGAAACGGGGTGTGACAGGCCCTC 1327
DB 2655 TGATGAGGGGATCAAGGCTCCGGGAGTGCCTCTGTGAAACAGGGTGTGACAGGCCCTC 2714
QY 1328 GTGTCACATCAGGACAGTTCCTGCTGAGTGTGACGCTCTGCTTCTGCTCATGCCAC 1387
DB 2715 GTGTGACATCTCCACAGCTGTATTCGAGTGTGACACCTGCTTCTGCTCGTCAACGCCAC 2774
QY 1388 CTGTGAAGGAGAAACAACAGCTGTGAGTGTAACTGGATTAAGAGTGTGACGGAATCAATG 1447
DB 2775 CTGTGACGAGAAACAACAGCTGTGAGTGTAACTGGATTAAGAGTGTGACGGAATCAATG 2834
QY 1448 CACAGTGTGTGATTTCTGCAACACAGGACAAACCGGGCTGTGCAAGGTGTGACGAGTGTCT 1507
DB 2835 CACAGTGTGTGATTTCTGCAACACAGGACAAACCGGGCTGTGCAAGGTGTGACGAGTGTCT 2894
QY 1508 CCAGAGGCGACAGAGTCTCTGCAAGTGTGCAAGGAGGATACAAAGGGGACGCGGACAG 1567
DB 2895 CCAGAGGCGACAGAGTCTCTGCAAGTGTGCAAGGAGGATACAAAGGGGACGCGGACAG 2954
QY 1568 CTGCAACAGAGATAGACCCCTGTGCAACCGGCTTAAACGAGGAGTGTCAACGAGACGCCAC 1627
DB 2955 CTGCAACAGAGATAGACCCCTGTGCAACCGGCTTAAACGAGGAGTGTCAACGAGACGCCAC 3014
QY 1628 CTGTAAGATGACAGGCGCGGACAGCAAGTGTGAGTGTAAAGTCACTATCTCGGAGA 1687
DB 3015 CTGCAAGATGACGCGGCGCGGACAGCAAGTGTGAGTGTAAAGTCACTATCTCGGAGA 3074
QY 1688 TGSGCTGAATGTGAGCGCGGACAGCTGCCATTTGACCGCTGCTTACAGGACAAATGGGCA 1747
DB 3075 CGAGTGTGAGTGTGAGCGCTGAGCAGCTGCGCTCGACCGTGTCTTACAGGACAAACGACA 3134
QY 1748 GTGCCATGAGAGCGCAAAATGTGTGCAACCTTCACTTCCAGGATACCACTGTTCGGGTGTT 1807
DB 3135 GTGCCACCCAGATGCCAGCTGTGCAACCTTCACTTCCAGGACAAACCGGACAAAGGCTG 3194
QY 1808 CCACTACGCTCCCGCTGCGGCGGAGTAAAGTGTGACCTTTGACAAAGCCAGAGAGGCGCTG 1867
DB 3195 CCACTACGCTCCCGCTGCGGCGGAGTAAAGTGTGACCTTTGACAAAGCCAGAGAGGCTG 3254
QY 1868 TGCCAAACGAGCTGGCAACCATGCAACCTTCAACACAGCTCTCTATGCCAGAGGCGCAA 1927
DB 3255 TGCCAAACGAGCTGGCAACCATGCAACCTTCAACACAGCTCTCTATGCCAGAGGCGCAA 3314
QY 1928 GTACCACTGTGTCTGAGAGGCTGGGTGTGAGAGACGGGCGGGTGTGCTTACCCACAGCCTT 1987
DB 3315 GTATCACTGTGTCTGCGCGGCTGGGTGTGAGAGTGGGCGGGTGTGCTTACCCGACTACGTA 3374
QY 1988 CGCTCCCAAGCTGTGGCTCTGTGTGGTGTGGATAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2047


```
Db 3375 TGCCCTCTCAGAGTGTGTGTCACAAACGTTGTTGGATCGTAGACTACGATCCAGGCGCAA 3434
Qy 2048 CAAGAGTGAATGTTGGAGTGTCTTCTGCTATCGATGAAGATGTGAATCTGCACCTGCAA 2107
Db 3435 CAAGAGTGAATGTTGGAGTGTCTTCTGCTATCGATGAAGATGTGAATCTGCACCTGCAA 3494
Qy 2108 GGTGGCTATGTGGAGATGGCTTCTCATGCACTGGGAACCTGCTCAGTCTCTGATGTC 2167
Db 3495 GGCAGGCTATGTGGAGATGGCTTCTCATGCACTGGGAACCTGCTCAGTCTCTGATGTC 3554
Qy 2168 CTTCCCTCCTCACTCAAACTTCTGACGGAAGTCTGGCTATTCACACAGCTCAGCTCG 2227
Db 3555 CTTCCCTCCTCACTCAAACTTCTGACGAGTGTCTGGCTTTTCCAGAGCTCAGCCCG 3614
Qy 2228 AGGCGGTGATTTCTAGACACTGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2287
Db 3615 AGGACAGGATTTTGAACAACCTGACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3674
Qy 2288 ACAGAACAGTGGGCTGGGGGAATGAGACCTTGTCTGGCGGGACATCGAGCACCACT 2347
Db 3675 ACAGAACAGTGGGCTACCGGGAATGAGAGCTGTCTGGCGGGACATGAGACCACT 3734
Qy 2348 CCCCAGTGTGACATGTTTCTCAATGACCTTGTCAATGACCACTGCTGCTGCTGCTGCTG 2407
Db 3735 CACTAATGTCAAGCTCTCTTTTACAATGACCTTGTCAATGACCTTGTCTGAGGACTAT 3794
Qy 2408 GGTGGGAAGCAAGCTGCTCATCTGCGAGCCAGGACCACTGCGAACCGAGGAGACGAG 2467
Db 3795 GCTGGGAAGCCACTGCTCAATACCTTCAGCCAGGACCACTGCTGCTGCTGCTGCTGCTG 3851
Qy 2468 GTTGTGTGATGAAGAGCACTTCTGAGTGGGACATCTTGTCTGCTCAATGGGATCATCA 2527
Db 3852 GTTGTGTGATGAAGATCACTTCTGAGTGGGACATCTGCTGCTCAATGGGATCATCA 3911
Qy 2528 TGTCAATTTCCAGGCTTTTAAAGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2587
Db 3912 TATTATTTCTGAACCTTTGAGAGCTCTTCCACGCGCAGCAACGCTGCTGCTGCTGCT 3971
Qy 2588 GGGAGCAGGATCTTCTTGTGATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2647
Db 3972 GGGACAGGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4031
Qy 2648 CTCTACTTTTGGATTAACCGGAGAACATCGCTTTCAGGATTTTGGATGCTGAGAGGA 2707
Db 4032 CTCTACTTTCGGCTTAAGCAGGACCACTGCTTTCAGGATTTTGA-TCAGAGAGGA 4090
Qy 2708 CATTAATGTTGCACTCTTGGGAGCAGCAGCCTGAGAAATCTCGAACCCCTTGTATGA 2767
Db 4091 CATTAATGTTGCTGCTTGGGAGCAGCAGCCTGAGAAATCTCGAACCCCTTGTATGA 4150
Qy 2768 GAGCACAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2827
Db 4151 GACTCAGCGCGGACACCCAGAGTCTCTGAGACCCCTTCAGACCCCTGAGAGCA 4210
Qy 2828 GCAGCTTGAGGCAATGACCCCTTGGAGACACTGTGAGGCGCTGAGGAGATGCGCAGC 2887
Db 4211 GGATCTGGGACAGCAGCCTCTGGGGCACTG---CGTCTGATGATGAGAGCAGC 4267
Qy 2888 CATCACTCACTGCGCAGCTGGGCAATCACTGTGAATCTCAGC---ACCAGTTGCCCTTT 2944
Db 4268 AAGCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4327
Qy 2945 AGGAACGTAAGTCTTTAAGCACTCAGAGCCATCTCTCTCTGCTGCTGCTGCTGCTG 3004
Db 4328 GTTTAAGATGACACTCAATAGCCAGCCATCTCACCCTTCTGCTGCTGCTGCTGCTG 4387
Qy 3005 GGTGTTTCTGTGGGTGAGAGATGTTG-CTGTGCCCCACCCAGTACAGCTTCTCTCT 3063
Db 4388 ATTGTGCCAGGCTAGGAGCCATGTTGCTGATACCTGCGGGGACCTCCACCTCTCT 4447
Qy 3064 GACCTTT 3071
Db 4448 GAGCCTAT 4455
```

```
RESULT 9
ABX13821
ID ABX13821 standard; cDNA; 4706 BP.
XX
AC ABX13821;
XX
DT 19-FEB-2003 (first entry)
XX
cDNA encoding rat 175kDa Hyaluronan receptor for endocytosis (HARE).
XX
Hyaluronan receptor for endocytosis; HARE; hyaluronan; HA; CD;
KW chondroitin molecule; chondroitin sulphate; CDS; targeting; metastasis;
KW tumour; gene therapy; rat; gene; ss.
XX
OS Rattus sp.
XX
Key Location/Qualifiers
CDS 1..4296
FT /*tag= a
FT /product= "HARE"
FT /note= "Hyaluronan receptor for endocytosis"
FT /partial
FT /note= "No start codon given"
XX
WO200286093-A2.
XX
31-OCT-2002.
XX
25-APR-2002; 2002WO-US013209.
XX
25-APR-2001; 2001US-00842930.
XX
25-APR-2001; 2001US-0286468P.
XX
(WEIG/) WEIGEL P H.
XX
(WEIG/) WEIGEL J A.
XX
Weigel PH, Weigel JA;
XX
WPI; 2003-093126/08.
XX
P-PSDB; ABG72498.
XX
Targeting compounds e.g. chemotherapeutic agent to cell of subject
expressing functional active hyaluronan receptor for endocytosis of HARE,
by using hyaluronan, HARE, and/or monoclonal antibody that binds to HARE
epitope.
XX
Example; Fig 2; 167pp; English.
XX
The invention describes a method of targeting a compound to a cell or
tissue of an individual expressing a functionally active hyaluronan (HA)
receptor for endocytosis (HARE) or a cell that does not express
functionally active HARE. The method involves using HA molecule, a
chondroitin molecule (CD), a chondroitin sulphate (CDS) molecule, and/or
a monoclonal antibody raised against a HA-binding domain of HARE. The
method is useful for targeting a compound, preferably a chemotherapeutic
agent or a radioisotope to cell of an individual, especially a human,
expressing HARE on its surface (e.g. gene therapy). Also described is a
method useful for preventing interaction between a cell having at least
one of a HA coat, a CD coat and CDS coat and a cell expressing HARE on
its surface. This second method is useful for preventing metastasis by
preventing interaction between tumour cells having HA, CD or CDS coat and
non-tumour cells expressing HARE on its surface. The invention also
describes a method useful for detecting the presence of HA, CD and CDS in
a biological fluid. This sequence encodes the rat 175kDa Hyaluronan
receptor for endocytosis (HARE)
XX
Sequence 4706 BP; 1168 A; 1273 C; 1230 G; 1035 T; 0 U; 0 Other;
```

```
Query Match 59.4%; Score 1936.8; DB 7; Length 4706;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 652; Indels 14; Gaps 6;
```


Db 3555 CTTCCCTCGCTCACAACCTCTCTGACAGAGTCTGCTCTTTTCCAGAGCTCAGCCG 3614
QY 2228 AGCCCGTGCAATTTCTAGAACACCTGACTGCTGTCCATCGCGGACACCTCTTTGTGCG 2287
Db 3615 AGACAGGCAATTTTGAACACCTGACTGCTGTCCATCGTGGCACCTGTTGTGCG 3674
QY 2288 ACAGAACAGTGGCTGGGGAGATGAGACCTTGTCTGGGGGAGCATCAGACACCACT 2347
Db 3675 ACAGAACAGTGGCTGGGGAGATGAGACCTTGTCTGGGGGAGCATCAGACACCACT 3734
QY 2348 CGCAATGTGAGCATGTTTCTTACATGACCTTGTCTCAATGGCACACCTCTGCAAGGAG 2407
Db 3735 CACTAATGTCAACGTCTCTCTTTTACAATGACCTTGTCTCAATGGTACCTTCTGAGGACTAT 3794
QY 2408 GGTGGGAGCAACTGCTCTATCTGCTGAGGACAGGACCACTCAACCGGAGGAGACCA 2467
Db 3795 GCTGGGAGCAACTGCTCTATCTGCTGAGGACAGGACCACTCAACCGGAGGAGACCA 3851
QY 2468 GTTGTGTGATGAGAGGCAATCTGCAAGTGGGACATCTTGTCTTCCATGGGATCATTC 2527
Db 3852 GTTGTGTGATGAGAGGCAATCTGCAAGTGGGACATCTTGTCTTCCATGGGATCATTC 3911
QY 2528 TGTCAATTTGAGGCTTTTAAAGCACCCCTGCTGCGCCGCTGACCTTGACCCACACTGGCT 2587
Db 3912 TATTATTTCTGAACCTTTGAGAGCTCTCTCCACGGGAGCAACGGCTGCGCCACTCTGGCT 3971
QY 2588 GGGAGCAGGAGTCTTCTTGGCATCATCTGCTGAGTGGGCTGTGGCTGTGGCTGTCTTA 2647
Db 3972 GGGGAGGAGTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4031
QY 2648 CTCCTACTTTCCGATAAAGCGGAGAACATCGGCTTCCAGCATTTTGTGCTGGAGAGA 2707
Db 4032 CTCCTACTTTCCGATAAAGCGGAGAACATCGGCTTCCAGCATTTTGTGCTGGAGAGA 4090
QY 2708 CATTATGTGTGAGTCTTTGGAAGCAGCAGCTGCTGAGATATCTCGAACCCCTGTGATGA 2767
Db 4091 CATTATGTGTGAGTCTTTGGAAGCAGCAGCTGCTGAGATATCTCGAACCCCTGTGATGA 4150
QY 2768 GAGCACAACCTCAGCTCCCGGAGACCTTCTTACGACCCCTTCCAGGACTCTGAGAACG 2827
Db 4151 GACCTCAGCGCGGACCCCGGAGACCTTCTTACGACCCCTTCCAGGACTCTGAGAACG 4210
QY 2828 GCAGCTTGAGGCAATGACCCCTTCCAGGACACTGTGAGGCTGACCGGAGATCCAGC 2887
Db 4211 GGATCTGAGACACGACCCCTTCTGGGGGCACTG---CGGTCTGACATGAGAGCCAGC 4267
QY 2888 CATCACTCACTGCCACCTGGGCAATCACTGTGAATTTCTCAGC---ACGAGTTCCTTTT 2944
Db 4268 AAGCAACACAGTCACGGTTCCAGGTAATTCACGACCCCTTCCAGGACTCTGATGATCAGTT 4327
QY 2945 AGGAACGTAAGTCTTTTAAAGCACTCAGAACGCTATCTCTGCTGCTGATCTGGG 3004
Db 4328 GTTTTAAAGATGACAACTCATAAGCAGGACCACTACCTTCTGCTGATCTGGG 4387
QY 3005 GGTGTTTCTGCTGGGAGAGATGTTG-CTGTCGCCACCCAGTACAGTCTCTCTCT 3063
Db 4388 ATTGTCGCCAGGCTAAGGAGCAATGTTGCTGCTGATACCTGCGGAGCCTCCACCTCTCT 4447
QY 3064 GACCTTTT 3071
Db 4448 GAGCCTAT 4455

RESULT 10
AAF87119
ID AAF87119 standard; DNA; 2483 BP.
XX
AC AAF87119;
XX
DT 26-MAR-2002 (first entry)
XX
DE NOV8 coding sequence.

XX
KW NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW cardiovascular; casein kinase II phosphorylation site; contraception;
KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW epidermal growth factor; cell development; apoptosis; cell adhesion;
KW inflammatory disorder; cell structure; motility; cancer; immune disorder;
KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.
XX
OS Unidentified.
XX
FH Key
FT CDS 387..2396
FT /tag= a
FT /product= "NOV8"
XX
PN WO200136638-A2.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-US031543.
XX
PR 19-NOV-1999; 99US-0166336P.
PR 29-NOV-1999; 99US-0167785P.
PR 08-MAR-2000; 2000US-0187844P.
PR 16-NOV-2000; 2000US-00715417.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Lichenstein H, Vernet C, Fernandes E;
XX WPI: 2001-648134/74.
XX P-PSDB; AAB83364.
XX
PT Novel human polypeptides and the nucleic acids that encode them useful
PT for preventing, diagnosing and treating e.g. cancer, inflammation and
PT immune disorders.
XX
PS Claim 9; Page 29-30; 141pp; English.
XX
CC This sequence encodes the NOV8 protein. The invention relates to the NOV1
CC -NOV16 proteins, and their coding sequences. The proteins have Cytostatic
CC ; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
CC activities. The sequences may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate NOVX expression. They
CC may be used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC activity of protein by expressing inactive proteins or to supplement the
CC patients own production of protein. They are used to produce NOVX
CC proteins, by inserting the nucleic acid into a cell and culturing it to
CC express the protein. The DNA may be used as DNA probes in assays to
CC detect and quantitate the presence of similar DNAs in samples, and which
CC patients may need restorative therapy. The NOVX protein may also be used
CC as antigens in the production of antibodies (Abs) against NOVX and in
CC assays to identify modulators of NOVX expression and activity. The anti-
CC NOVX Abs and antagonist are used to down regulate expression and
CC activity. The anti-NOVX Abs are used for detecting the presence of NOVX
CC in samples. Disorders that may be prevented, diagnosed and/or treated
CC vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
CC NOV13-16 have casein kinase II phosphorylation sites characteristic of
CC serine/threonine kinases, and are used to treat kinase-related disorders
CC (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
CC -like super family and are involved in, e.g. regulation of cell
CC development, apoptosis, cell adhesion, growth migration, cell structure
CC and motility and protein management, and are used to treat cancers,
CC inflammatory disorders, immune disorders and cellular adhesion disorders.
CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
CC treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
CC syndrome and marfan syndrome
XX

SQ	Sequence	2483 BP; 619 A; 648 C; 639 G; 577 T; 0 U; 0 Other;	
	Query Match	54.6%; Score 1778.6; DB 4; Length 2483;	
	Best Local Similarity	98.9%; Pred. No. 0;	
	Matches 1790; Conservative	0; Mismatches 19; Indels 0; Gaps 0;	
Qy	122	AGCCTTTCATGAGGAGCTCGGTTTAAAGATCGGACAAATACCGTTTAAATGCCCGCAGGT	181
Db	569	ACCCCTCTGATCTTCTGTACAGGTTAAAGATCGGACAAATACCGTTTAAATGCCCGCAGGT	628
Qy	182	TCTTCGGTACCATGTCGCTGCCACAGCTGCTCTGGAACACCTGAAATGATCTC	241
Db	629	TCTTCGGTACCATGTCGCTGCCACAGCTGCTCTGGAACACCTGAAATGATCTC	688
Qy	242	AAATGCTACTTCCCTCCAAAGAGAGCCAAATAGTCAATCTCCGTCCTCGAGACACGGTGA	301
Db	689	AAATGCTACTTCCCTCCAAAGAGAGCCAAATAGTCAATCTCCGTCCTCGAGACACGGTGA	748
Qy	302	TATAAATTAAGGCTTAAGATCATATCCAGTGATATCATCATCTAGTACTTAATGGGATTTGTTCA	361
Db	749	TATAAATTAAGGCTTAAGATCATATCCAGTGATATCATCATCTAGTACTTAATGGGATTTGTTCA	808
Qy	362	TATCATAGACAATTTGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGG	421
Db	809	TATCATAGACAATTTGCTATCTCCCAAAATTTGCTTATCACTCCCAAGACAACTCTGG	868
Qy	422	AAGAATTCGCAAAATTTTACGACTTTGGCAACAACAATGGGTGATCAATCAAAATTTAGCAA	481
Db	869	AAGAATTCGCAAAATTTTACGACTTTGGCAACAACAATGGGTGATCAATCAAAATTTAGCAA	928
Qy	482	CTTAATACAGGACTCAGGTTTGTGAGTGTATCATCCGATCCCATCCACACCCAGTCAAC	541
Db	929	CTTAATACAGGACTCAGGTTTGTGAGTGTATCATCCGATCCCATCCACACCCAGTCAAC	988
Qy	542	TCTCTTCTGSCCACCAGCAAGCCCTCCATGCTTACCTGCTGCTGGAACAACAGGACTTCCT	601
Db	989	TCTCTTCTGSCCACCAGCAAGCCCTCCATGCTTACCTGCTGCTGGAACAACAGGACTTCCT	1048
Qy	602	GTTTCAACCAAGACAACAAGGCAAGCTGAAGAGTATTTGAAGTTTCTATGTGATAGAGA	661
Db	1049	GTTTCAACCAAGACAACAAGGCAAGCTGAAGAGTATTTGAAGTTTCTATGTGATAGAGA	1108
Qy	662	TGCCAAGGTTTACGTTGATCTTCCACATCCACTGCTGCTGGAAGACCCCTGCAAGGTC	721
Db	1109	TGCCAAGGTTTACGTTGATCTTCCACATCCACTGCTGCTGGAAGACCCCTGCAAGGTC	1168
Qy	722	AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTGAAATGGCCA	781
Db	1169	AGAGCTGAGTGTGAAATGTGGAGCTGGCAGGACATCGGTGACCTCTTCTGAAATGGCCA	1228
Qy	782	AACCTGCAGAAATTGTGCAGCGGAGCTCTTTGTTGACCTGGGTGTGCGCTTACCGCATTTGA	841
Db	1229	AACCTGCAGAAATTGTGCAGCGGAGCTCTTTGTTGACCTGGGTGTGCGCTTACCGCATTTGA	1288
Qy	842	CTGTCTGCTGATTCATCCACCTCGGGGCGCTGTGACACCTTTTACTACTTTTCGATGC	901
Db	1289	CTGTCTGCTGATTCATCCACCTCGGGGCGCTGTGACACCTTTTACTACTTTTCGATGC	1348
Qy	902	CTCGGGGAGTGTGGGAGCTGTGTCAATCTCCAGCTGCTGCAAGGGCTACTTTCCGGCGAGA	961
Db	1349	CTCGGGGAGTGTGGGAGCTGTGTCAATCTCCAGCTGCTGCAAGGGCTACTTTCCGGCGAGA	1408
Qy	962	GGGTGTGAAGCAGAAAGTGTCTCTCAACCTGCTCCCTCAAGAGGACCTGGAAGCTGCG	1021
Db	1409	GGGTGTGAAGCAGAAAGTGTCTCTCAACCTGCTCCCTCAAGAGGACCTGGAAGCTGCG	1468
Qy	1022	GGAGGGTGCAGCTGGTGATACAGATCCCCAGGCTGTGCAAGGGCTACTTTCCGGCGAGA	1081
Db	1469	GGAGGGTGCAGCTGGTGATACAGATCCCCAGGCTGTGCAAGGGCTACTTTCCGGCGAGA	1528
Qy	1082	CTGTGAGGCTTCCCTCGAGGACCAAGATGCCCCGTTAATACCGGGGTGTCTGCTTGA	1141
Db	1529	CTGTGAGGCTTCCCTCGAGGACCAAGATGCCCCGTTAATACCGGGGTGTCTGCTTGA	1588

Qy	1142	TCAGTACTCGGCCACCGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA	1201
Db	1589	TCAGTACTCGGCCACCGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGGCGTGTGA	1648
Qy	1202	GATGTGTGGCGGGGAGATTTCGGGCGCTGATTTGCTGCCCTGTGGCTCTCAGACCCAGG	1261
Db	1649	GATGTGTGGCGGGGAGATTTCGGGCGCTGATTTGCTGCCCTGTGGCTCTCAGACCCAGG	1708
Qy	1262	ACAGTGGCATGATGCGATCAACGGGCTCGGGCAGTGCCTCTGTGAAAACGGGGTGACAGG	1321
Db	1709	ACAGTGGCATGATGCGATCAACGGGCTCGGGCAGTGCCTCTGTGAAAACGGGGTGACAGG	1768
Qy	1322	CCCTCTGCTGTGACATCTCAGGCGAGTTTTCCTCTGCAAGTGTGAGCCCTCTTGTCTCTCA	1381
Db	1769	CCCTCTGCTGTGACATCTCAGGCGAGTTTTCCTCTGCAAGTGTGAGCCCTCTTGTCTCTCA	1828
Qy	1382	TGCCACCTGTGAAGAGAAACAACGCTGTGAGTGTAACTGGGATTAAGAGTGAACGGAAT	1441
Db	1829	TGCCACCTGTGAAGAGAAACAACGCTGTGAGTGTAACTGGGATTAAGAGTGAACGGAAT	1888
Qy	1442	CACATGCAAGTTTGTGATTTCTGCAAAACAGGACAAACGGGGCTGTGAAAAGTGGCCAG	1501
Db	1889	CACATGCAAGTTTGTGATTTCTGCAAAACAGGACAAACGGGGCTGTGAAAAGTGGCCAG	1948
Qy	1502	ATGCTCCCAGAGGCGACGAAGTCTCTGCAAGTGTGCAAGAGGATACAAAGGGGACGG	1561
Db	1949	ATGCTCCCAGAGGCGACGAAGTCTCTGCAAGTGTGCAAGAGGATACAAAGGGGACGG	2008
Qy	1562	GCACAGCTGCACAGAGATAGACCCCTGTGCACAGCGCTTAACGGAGGCTGTCAAGGCA	1621
Db	2009	GCACAGCTGCACAGAGATAGACCCCTGTGCACAGCGCTTAACGGAGGCTGTCAAGGCA	2068
Qy	1622	CGCCACCTGTAAAGTGAAGAGTGTGAGCGCGGCAAGCAAGTGTGAGTGTAAAGTCACTATGT	1681
Db	2069	CGCCACCTGTAAAGTGAAGAGTGTGAGCGCGGCAAGCAAGTGTGAGTGTAAAGTCACTATGT	2128
Qy	1682	CGGAGATGGGCTGAACTGTGAGCGGAGCAGTGTGCCCTTGAAGTGTCTTACAGGACAA	1741
Db	2129	CGGAGATGGGCTGAACTGTGAGCGGAGCAGTGTGCCCTTGAAGTGTCTTACAGGACAA	2188
Qy	1742	TGGGAGTGGCATGCGAGACCCCAATGTGTGACCTTCCAGTACCACTCTGTTGG	1801
Db	2189	TGGGAGTGGCATGCGAGACCCCAATGTGTGACCTTCCAGTACCACTCTGTTGG	2248
Qy	1802	GGTGTTCATCTACGCTCCCACTGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA	1861
Db	2249	GGTGTTCATCTACGCTCCCACTGGCCAGTATAAGCTGACCTTTGACAAAGCCAGAGA	2308
Qy	1862	GGCTGTGCCAACGAAGCTGCGACCAATGGCAACTTCAACAGCTCTCTTATGCCAGAA	1921
Db	2309	GGCTGTGCCAACGAAGCTGCGACCAATGGCAACTTCAACAGCTCTCTTATGCCAGAA	2368
Qy	1922	GGCCAAAGTA 1930	
Db	2369	GACCTGGTA 2377	

RESULT 11
AAP87117
ID AAP87117 standard; DNA; 2512 BP.
XX
AC AC
XX XX
DT 26-MAR-2002 (first entry)
XX
DE NOV6 coding sequence.
XX
KW NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW epidermal growth factor; cell development; apoptosis; cell adhesion;
KW growth migration; cell structure; motility; cancer; immune disorder;

Db 1678 GATGTGCTGGCGGAGATTGGCCCTGATTTGCTGCCCTGTGGCTGCTCAGACCAGG 1737
Qy 1262 ACAGTGCCATGATGTCATCAGCGGCTCCGGGAGTGCCTCTGTGAAACGGGGTGGACAGG 1321
Db 1738 ACAGTGCCATGATGTCATCAGCGGCTCCGGGAGTGCCTCTGTGAAACGGGGTGGACAGG 1797
Qy 1322 CCCCTGCTGACATCAGGCGATTTTGCCTGTCAGTGTGTAGCCCTCTTGTTCGCTCA 1381
Db 1798 CCCCTGCTGACATCAGGCGATTTTGTCTGCACTGTGTAGCCCTCTTGTTCGCTCA 1857
Qy 1382 TGCCACCTGTAAGGAGAACACACCTGTGTAGTGAACCTGGAATTATGAAGTGAACGGAAT 1441
Db 1858 TGCCACCTGTAAGGAGAACACACCTGTGTAGTGAACCTGGAATTATGAAGTGAACGGAAT 1917
Qy 1442 CACATGTCACAGTGTGTGATTTTCTGCAAAACAGGACAAACGGGGCTGTGAAAGTGGCCAG 1501
Db 1918 CACATGTCACAGTGTGTGATTTTCTGCAAAACAGGACAAACGGGGCTGTGAAAGTGGCCAG 1977
Qy 1502 ATGCTCCAGAGGCGACGAGGCTCTCTGTCAGCTGCCAGAGGATCAAGAGGAGCGG 1561
Db 1978 ATGCTCCAGAGGCGACGAGGCTCTCTGTCAGCTGCCAGAGGATCAAGAGGAGCGG 2037
Qy 1562 GCACAGCTGCACAGAGATAGACCCCTGTGCACAGCGCCCTTAACGGAGGGTGTCAAGGACA 1621
Db 2038 GCACAGCTGCACAGAGATAGACCCCTGTGCACAGCGCCCTTAACGGAGGGTGTCAAGGACA 2097
Qy 1622 CGCCACCTGTAAGATGACAGCGCCCGGCAAGCACAAAGTGTGTGATTAAGTCACTATGT 1681
Db 2098 CGCCACCTGTAAGATGACAGCGCCCGGCAAGCACAAAGTGTGTGATTAAGTCACTATGT 2157
Qy 1682 CGGAGATGGCTGAATGTGACCGGAGCAGCTGCCATTGACCGCTGCTTACAGGACAA 1741
Db 2158 CGGAGATGGCTGAATGTGACCGGAGCAGCTGCCATTGACCGCTGCTTACAGGACAA 2217
Qy 1742 TGGGAGTCCCATGACAGCGCCAAATGTGTGACCTCCACTTCCAGATACCACTGTGTGG 1801
Db 2218 TGGGAGTCCCATGACAGCGCCAAATGTGTGACCTCCACTTCCAGATACCACTGTGTGG 2277
Qy 1802 GGTGTTCATATACGCTCCCACTGGGCGAGTATAGCTGACCTTTGCAAAAGCCAGAGA 1861
Db 2278 GGTGTTCATATACGCTCCCACTGGGCGAGTATAGCTGACCTTTGCAAAAGCCAGAGA 2337
Qy 1862 GGCCTGTGCAAGAGCTGGACCATGCACTCAACCACTCTCTATGCCAGAA 1921
Db 2338 GGCCTGTGCAAGAGCTGGACCATGCACTCAACCACTCTCTATGCCAGAA 2397
Qy 1922 GGCCTAGTA 1930
Db 2398 GACCTGTA 2406

RESULT 12

AAF87113
ID AAF87113 standard; DNA; 2011 BP.
XX AAF87113;
AC AAF87113;
DT 26-MAR-2002 (first entry)
XX NOV2 coding sequence.
XX NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
KW cardiovascular; casein kinase II phosphorylation site; contraception;
KW serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
KW epidermal growth factor; cell development; apoptosis; cell adhesion;
KW growth migration; cell structure; motility; cancer; immune disorder;
KW inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
KW cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
KW therapy; NOV1; NOV2; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
KW NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.
XX Unidentified.

XX Key Location/Qualifiers
FH 877..1824
FT /*tag= a
FT /product= "NOV2"
PN WO200136638-A2.
XX 25-MAY-2001.
XX 17-NOV-2000; 2000WO-US031543.
XX 19-NOV-1999; 99US-0166336P.
PR 29-NOV-1999; 99US-0167785P.
PR 08-MAR-2000; 2000US-0187844P.
PR 16-NOV-2000; 2000US-00715417.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Lichenstein H, Vernet C, Fernandes E;
DR WPI: 2001-648134/74.
DR P-PSDB; AAB83358.
XX Novel human polypeptides and the nucleic acids that encode them useful
for preventing, diagnosing and treating e.g. cancer, inflammation and
immune disorders.
XX Claim 9; Page 10-13; 141pp; English.
XX This sequence encodes the NOV2 protein. The invention relates to the NOV1
-NOV16 proteins, and their coding sequences. The proteins have Cytostatic
; contraceptive; antiinflammatory; immunomodulatory; and cardiovascular
activities. The sequences may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate NOVX expression. They
may be used to treat disorders associated with decreased expression by
rectifying mutations or deletions in a patient's genome that affect the
activity of protein by expressing inactive proteins or to supplement the
patients own production of protein. They are used to produce NOVX
proteins, by inserting the nucleic acid into a cell and culturing it to
express the protein. The DNA may be used as DNA probes in assays to
detect and quantitate the presence of similar DNAs in samples, and which
patients may need restorative therapy. The NOVX protein may also be used
as antigens in the production of antibodies (Abs) against NOVX and in
assays to identify modulators of NOVX expression and activity. The anti-
NOVX Abs and antagonist are used to down regulate expression and
activity. The anti-NOVX Abs are used for detecting the presence of NOVX
in samples. Disorders that may be prevented, diagnosed and/or treated
vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and
NOV13-16 have casein kinase II phosphorylation sites characteristic of
serine/threonine kinases, and are used to treat kinase-related disorders
(e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).
CC NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)
-like super family and are involved in, e.g. regulation of cell
development, apoptosis, cell adhesion, growth migration, cell structure
and motility and protein management, and are used to treat cancers,
inflammatory disorders, immune disorders and cellular adhesion disorders.
CC NOV6-10 are homologous to EGF-like fibrillin proteins and are used to
treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT
syndrome and marfan syndrome
XX Sequence 2011 BP; 458 A; 551 C; 583 G; 419 T; 0 U; 0 Other;
Query Match 52.8%; Score 1721.2; DB 4; Length 2011;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 1843; Conservative 0; Mismatches 8; Indels 107; Gaps 1;
Qy 487 TACAGAGCTCAGCTTTGCTGAGTGTATCCAGATCCATCCACCCAGTCACTCTCT 546
Db 161 TTCAGGACTCAGCTTTGCTGAGTGTATCCAGATCCATCCACCCAGTCACTCTCT 220
Qy 547 TCTGGCCACCGACCGCCCTCCATCGCTACTCTGCTGAACACAGAGCTTCTCTGTCA 606

Db 221 TCTGGCCACCGACCAAGCCCTCCATGCTGCTGAAACAACAGGACTTCCTGTTCA 280
QY 607 ACCAAGACAACAGGACAGCTGAAGAGTATTTGAAGTTTTCATGTGATACGAGATGCCA 666
Db 281 ACCAAGACAACAGGACAGCTGAAGAGTATTTGAAGTTTTCATGTGATACGAGATGCCA 340
QY 667 AGGTTTGTAGCTGTGATCTTCCACATCCACTGCTGGAGACCTTGAAGTTTCAGAGC 726
Db 341 AGGTTTGTAGCTGTGATCTTCCACATCCACTGCTGGAGACCTTGAAGTTTCAGAGC 400
QY 727 TGAGTGTGAATGTGGAGCTGGCAGGGAATCGGTGACCTCTTTCTGAATGCCAAACCT 786
Db 401 TGAGTGTGAATGTGGAGCTGGCAGGGAATCGGTGACCTCTTTCTGAATGCCAAACCT 460
QY 787 GCAGAAATGTGAGCGGGAGCTCTTGTGTTGACTGGGTGCGCTACCGCATTTGACTGTC 846
Db 461 GCAGAAATGTGAGCGGGAGCTCTTGTGTTGACTGGGTGCGCTACCGCATTTGACTGTC 520
QY 847 TGCTGATTGATCCACCTCGGGGGCGCTGTGACACCTTTTACTACTTTTCGATGCTCGG 906
Db 521 TGCTGATTGATCCACCTCGGGGGCGCTGTGACACCTTTTACTACTTTTCGATGCTCGG 580
QY 907 GGGAGTGTGGAGCTGTGTCAATACTCCAGCTGCCAAGTGGAGTAAACCAAGGGTG 966
Db 581 GGGAGTGTGGAGCTGTGTCAATACTCCAGCTGCCAAGTGGAGTAAACCAAGGGTG 640
QY 967 TGAAGCAGAGTGTCTTAACACTGCGCTTCAAGAGGAACCTGGAAGCTGCCGGGAGC 1026
Db 641 TGAAGCAGAGTGTCTTAACACTGCGCTTCAAGAGGAACCTGGAAGCTGCCGGGAGC 700
QY 1027 GGTGAGCTGTGTATACAGATCCCCAGCTGCTGCAAGGGTACTTCGGGCGAGACTGTC 1086
Db 701 GATGAGCTGTGTATACAGATCCCCAGCTGCTGCAAGGGTACTTCGGGCGAGACTGTC 760
QY 1087 AGGCTGCTCCCTGGAGGACAGATGCCCGTGTATAACCGGGGTGTCTGCCCTTGATCAGT 1146
Db 761 AGGCTGCTCCCTGGAGGACAGATGCCCGTGTATAACCGGGGTGTCTGCCCTTGATCAGT 820
QY 1147 ACTGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGCGTGTGAGATGT 1206
Db 821 ACTGGCCACCGGAGAGTGTAAATGCAACACCGGCTTCAATGGGACGCGTGTGAGATGT 880
QY 1207 GCTGCGCGGGGAGATTTCGGGCTGATTCTGCTGCCCTGTGTAACCGGGGTGTCTGATCAGT 1266
Db 881 GCTGCGCGGGGAGATTTCGGGCTGATTCTGCTGCCCTGTGTAACCGGGGTGTCTGATCAGT 940
QY 1267 GCGATGATGCGATCAACCGGCTCCGGGAGTGCCTCTGTGAAACCGGGGTGGACAGCCCT 1326
Db 941 GCGATGATGCGATCAACCGGCTCCGGGAGTGCCTCTGTGAAACCGGGGTGGACAGCCCT 1000
QY 1327 CGTGTGACACTCAGGCACTTTTGGCTGCGAGTGTGTACGCTCTTGTGTTCTGCTCATGCCA 1386
Db 1001 CGTGTGACACTCAGGCACTTTTGGCTGCGAGTGTGTACGCTCTTGTGTTCTGCTCATGCCA 1060
QY 1387 CCTTAAGGAGAACAAACAGTGTGAGTGTACCTGGATTATGAAGGTACCGGATACAT 1446
Db 1061 CCTTAAGGAGAACAAACAGTGTGAGTGTACCTGGATTATGAAGGTACCGGATACAT 1120
QY 1447 GCACAGTGTGGATTTCTGCAAAACAGGACCAACGGGGCTGTGCAAAAGGTGCCAGATGCT 1506
Db 1121 GCACAGTGTGGATTTCTGCAAAACAGGACCAACGGGGCTGTGCAAAAGGTGCCAGATGCT 1180
QY 1507 CCCAGAGGGGACGAAGTCTCTGACAGTCCAGAGAGGGATACAAAGGGGACGGGCACA 1566
Db 1181 CCCAGAGGGGACGAAGTCTCTGACAGTCCAGAGAGGGATACAAAGGGGACGGGCACA 1240
QY 1567 GCTCAGAGATAGACCCCTGTGACAGCGCTTAAAGGGGTGTACAGAGCACGCCA 1626
Db 1241 GCTCAGAGATAGACCCCTGTGACAGCGCTTAAAGGGGTGTACAGAGCACGCCA 1300
QY 1627 CCTGTAAGATGACAGGGCCCGGGCAGCAGCAAGTGTGAGTGTAAAGTCACTATCTCGGAG 1686
Db 1301 CCTGTAAGATGACAGGGCCCGGGCAGCAGCAAGTGTGAGTGTAAAGTCACTATCTCGGAG 1360

QY 1687 ATGGGCTGAACCTGTGAGCCGAGCAGCTGCCATTGACCGCTGCTTACAGGACAATGGGC 1746
Db 1361 ATGGGCTGAACCTGTGAGCCGAGCAGCTGCCATTGACCGCTGCTTACAGGACAATGGGC 1420
QY 1747 AGTGCCATGCGAGACGCCAAATGTGCGACCTCCACTTCCAGGATACCACTGTGGGGTGT 1806
Db 1421 AGTGCCATGCGAGACGCCAAATGTGCGACCTCCACTTCCAGGATACCACTGTGGGGTGT 1480
QY 1807 TCCATCTAGCTCCCACTGGCCGAGTATAGCTGACCTTTGCAAAAGCAGAGGGCT 1866
Db 1481 TCCATCTAGCTCCCACTGGCCGAGTATAGCTGACCTTTGCAAAAGCAGAGGGCT 1540
QY 1867 GTGCCAACGAAGCTCGACACCTGACCACTCAAAACAGCTCTCTATGCCCAGAGGCCA 1926
Db 1541 GTGCCAACGAAGCTCGACACCTGACCACTCAAAACAGCTCTCTATGCCCAGAGGCCA 1600
QY 1927 AGTACCACTGTGCTCAGAGGCTGGTGAGACCGGGCGGGTTCCTACCCACAGCT 1986
Db 1601 AGTACCACTGTGCTCAGAGGCTGGTGAGACCGGGCGGGTTCCTACCCACAGCT 1660
QY 1987 TCGCCTCCAGAACTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2046
Db 1661 TCGCCTCCAGAACTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1720
QY 2047 ACAAGAGTGAATGTGGAGTGTCTTCTGCTATCGGATGAAAGATGTGAACCTGCACTGCA 2106
Db 1721 ACAAGAGTGAATGTGGAGTGTCTTCTGCTATCGGATGAAA----- 1761
QY 2107 AGGTGGGTATGTGGAGATGGTCTCATGCAAGTGGGAACCTGCTGCAAGTCTCTGATGT 2166
Db 1762 ----- 1761
QY 2167 CTTTCCCTCACTCACAACTTCTGACGGAAGTGTGCTGCTTATTCACAGCTCAGCTC 2226
Db 1762 -----GGAAGTGTGCTGCTTATTCACAGCTCAGCTC 1793
QY 2227 GAGCCGTGCAATTTCTAGAACACCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2286
Db 1794 GAGCCGTGCAATTTCTAGAACACCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1853
QY 2287 CACAGACAGTGGCTGGGGAGAAATGAGACTTGTCTGCGCGGGACATCAGACCAACC 2346
Db 1854 CACAGACAGTGGCTGGGGAGAAATGAGACTTGTCTGCGCGGGACATCAGACCAACC 1913
QY 2347 TCGCCAAATGTCAGCATGTTTCTTACAAATGACCTTGTCAATGCGACCACTGCAACGA 2406
Db 1914 TCGCCAAATGTCAGCATGTTTCTTACAAATGACCTTGTCAATGCGACCACTGCAACGA 1973
QY 2407 GGTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGAC 2444
Db 1974 GGCTGGGAAGCAAGCTGCTCATCACTGCCAGCCAGGAC 2011

RESULT 13

AAF87114

ID AAF87114 standard; DNA; 1804 BP.

XX AC AAF87114;

XX AC AAF87114;

DT 26-MAR-2002 (first entry)

DE XX

XX XX

XX XX

NOV; Cytostatic; contraceptive; antiinflammatory; immunomodulatory;
cardiovascular; casein kinase II phosphorylation site; contraception;
serine/threonine kinase; Peutz-Jeghers syndrome; cellular proliferation;
epidermal growth factor; cell development; apoptosis; cell adhesion;
growth migration; cell structure; motility; cancer; immune disorder;
inflammatory disorder; cellular adhesion disorder; long-QT syndrome;
cardiovascular disease; hypertrophic cardiomyopathy; marfan syndrome;
therapy; NOV1; NOV3; NOV4; NOV5; NOV6; NOV7; NOV8; NOV9; NOV10;
NOV11; NOV12; NOV13; NOV14; NOV15; NOV16; ds.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX FT 877. .1611

XX FT /*tag= a

XX FT /product= "NOV3"

XX PN WO200136638-A2.

XX PD 25-MAY-2001.

XX 17-NOV-2000; 2000WO-US031543.

XX 19-NOV-1999; 99US-0166336P.

XX 29-NOV-1999; 99US-0167785P.

XX 08-MAR-2000; 2000US-0187844P.

XX 16-NOV-2000; 2000US-00715417.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Lichenstein H, Vernet C, Fernandes E;

XX WPI: 2001-648134/74.

XX P-PSDB; AAB93359.

XX Novel human polypeptides and the nucleic acids that encode them useful

XX for preventing, diagnosing and treating e.g. cancer, inflammation and

XX immune disorders.

XX Claim 9; Page 14-17; 141pp; English.

XX This sequence encodes the NOV3 protein. The invention relates to the NOV1

XX ; NOV16 proteins, and their coding sequences. The proteins have Cytostatic

XX ; contractive; antiinflammatory; immunomodulatory; and cardiovascular

XX activities. The sequences may be used in the prevention, diagnosis and

XX treatment of diseases associated with inappropriate NOVX expression. They

XX may be used to treat disorders associated with decreased expression by

XX rectifying mutations or deletions in a patient's genome that affect the

XX activity of protein by expressing inactive proteins or to supplement the

XX patients own production of protein. They are used to produce NOVX

XX proteins, by inserting the nucleic acid into a cell and culturing it to

XX express the protein. The DNA may be used as DNA probes in assays to

XX detect and quantitate the presence of similar DNAs in samples, and which

XX patients may need restorative therapy. The NOVX protein may also be used

XX as antigens in the production of antibodies (Abs) against NOVX and in

XX assays to identify modulators of NOVX expression and activity. The anti-

XX NOVX Abs and antagonist are used to down regulate expression and

XX activity. The anti-NOVX Abs are used for detecting the presence of NOVX

XX in samples. Disorders that may be prevented, diagnosed and/or treated

XX vary depending on the NOVX protein. NOV1, NOV3, NOV5, NOV7, NOV9-11 and

XX NOV13-16 have casein kinase II phosphorylation sites characteristic of

XX serine/threonine kinases, and are used to treat kinase-related disorders

XX (e.g. Peutz-Jeghers syndrome, cellular proliferation and contraception).

XX NOV2-3, NOV6 and NOV8 are homologous to the epidermal growth factor (EGF)

XX -like super family and are involved in, e.g. regulation of cell

XX development, apoptosis, cell adhesion, growth migration, cell structure

XX and motility and protein management, and are used to treat cancers,

XX inflammatory disorders, immune disorders and cellular adhesion disorders.

XX NOV6-10 are homologous to EGF-like fibrillin proteins and are used to

XX treat cardiovascular disease e.g. hypertrophic cardiomyopathy, long-QT

XX syndrome and marfan syndrome

XX SQ Sequence 1804 BP; 426 A; 487 C; 521 G; 370 T; 0 U; 0 Other;

XX Query Match 43.9%; Score 1429.6; DB 4; Length 1804;

XX Best Local Similarity 99.7%; Pred. No. 0;

XX Matches 1432; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 487 TAGAGACTCAGGTTTCTGAGTGTATCCAGATCCATCCATCCACCCAGTCACTCTCT 546

Db 161 TTCAGGACTCAGGTTTCTGAGTGTATCCAGATCCATCCATCCACCCAGTCACTCTCT 220

QY 547 TCTGGCCACCGCAACAGCCCTCCATGCCCTACCTGCTGAAACAAACAGAGCTTCTGTTCA 606

Db 221 TCTGGCCACCGCAACAGCCCTCCATGCCCTACCTGCTGAAACAAACAGAGCTTCTGTTCA 280

QY 607 ACCAAGACAACAAGCAAGCTGAAGAGTATTTGAAGTTTTCATGTGATAGAGATGCCA 666

Db 281 ACCAAGACAACAAGCAAGCTGAAGAGTATTTGAAGTTTTCATGTGATAGAGATGCCA 340

QY 667 AGGTTTATGCTGTGGATCTTCCCATCATCTGCTGGAAGACCCCTGCAAGGTTTCAGAGC 726

Db 341 AGGTTTATGCTGTGGATCTTCCCATCATCTGCTGGAAGACCCCTGCAAGGTTTCAGAGC 400

QY 727 TGAGTGTCAAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGCCAAACCT 786

Db 401 TGAGTGTCAAAATGTGGAGCTGGCAGGGACATCGGTGACCTCTTTCTGAATGCCAAACCT 460

QY 787 GCAGAAATGTGACGGGAGCTCTTTGTTGACCTGGGTGTGCTTACGGCATTTGACTGTC 846

Db 461 GCAGAAATGTGACGGGAGCTCTTTGTTGACCTGGGTGTGCTTACGGCATTTGACTGTC 520

QY 847 TGCTGATTGATCCACCCCTGGGGGCGCTGTGACACCTTTTACTACTTTTCGATGCTCGG 906

Db 521 TGCTGATTGATCCACCCCTGGGGGCGCTGTGACACCTTTTACTACTTTTCGATGCTCGG 580

QY 907 GGGAGTGTGGAGCTGTCTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG 966

Db 581 GGGAGTGTGGAGCTGTCTCAATATCTCCAGCTGCCCAAGGTGGAGTAAACCAAGGGTG 640

QY 967 TGAAGCAGAAGTGTCTTACAACCTGCCCTTCAAGAGNAACCTGGAAGCTCCCGGAGC 1026

Db 641 TGAAGCAGAAGTGTCTTACAACCTGCCCTTCAAGAGNAACCTGGAAGCTCCCGGAGC 700

QY 1027 GGTGACGCTGTGTATACAGATCCCGAGTGTCTGCAAGGGCTACTTTCGGGGCAGAGTCTC 1086

Db 701 GATGACGCTGTGTATACAGATCCCGAGTGTCTGCAAGGGCTACTTTCGGGGCAGAGTCTC 760

QY 1087 AGGCTTCCCTCGAGGACAGATGCCCGGTGTAATAACCGGGGTGTCTGCTTGTATCAGT 1146

Db 761 AGGCTTCCCTCGAGGACAGATGCCCGGTGTAATAACCGGGGTGTCTGCTTGTATCAGT 820

QY 1147 ACTGGCCACCGGAGAGTGTAAATGCACACCGGCTTCAATGGGACGGGTGTGAGATGT 1206

Db 821 ACTCGGCCACCGGAGAGTGTAAATGCACACCGGCTTCAATGGGACGGGTGTGAGATGT 880

QY 1207 GCTGGCGGGGAGATTCGGGGCTGATTGTCTGCTGTGCTGCTGCTCAGACCCAGCAGT 1266

Db 881 GCTGGCGGGGAGATTCGGGGCTGATTGTCTGCTGTGCTGCTCAGACCCAGCAGT 940

QY 1267 GCGATGATGCGCATCAGGGCTCCGGGCACTGCTCTGTGAAACCGGGGTGACAGGCCCT 1326

Db 941 GCGATGATGCGCATCAGGGCTCCGGGCACTGCTCTGTGAAACCGGGGTGACAGGCCCT 1000

QY 1327 CGTGTGACACTCAGGAGTTTCCCTGCGAGTGTACGCTCTCTTCTGTCTCATGCCA 1386

Db 1001 CGTGTGACACTCAGGAGTTTCTGTCGAGTGTGTACGCTCTCTTCTGTCTCATGCCA 1060

QY 1387 CCTGTAAAGAGAAACAACAGCTGTGAGTGTAAACCTGGAATTATGAAGGTGACGAATCACAT 1446

Db 1061 CCTGTAAAGAGAAACAACAGCTGTGAGTGTAAACCTGGAATTATGAAGGTGACGAATCACAT 1120

QY 1447 GCAAGTGTGGATTTCTGCAAAACAGGACAAACCGGGGCTGTGCAAGGTTGCCAGATGCT 1506

Db 1121 GCAAGTGTGGATTTCTGCAAAACAGGACAAACCGGGGCTGTGCAAGGTTGCCAGATGCT 1180

QY 1507 CCAGAGGGGCAAGAGTCTCTGCGAGTGCAGAGGGGATACAAAGGGGACGGGCACA 1566

Db 1181 CCAGAGGGGCAAGAGTCTCTGCGAGTGCAGAGGGGATACAAAGGGGACGGGCACA 1240

QY 1567 GCTGCAACAGAGATAGACCCCTGTGCGAGACGGCCCTTAAACGGAGGTGTACAGACGCCA 1626

Db 1241 GCTGCAACAGAGATAGACCCCTGTGCGAGACGGCCCTTAAACGGAGGTGTACAGACGCCA 1300

QY 1627 CCTGTAAAGTGAACAGGCCCGGGCAACAGTGTGTAGTGTAAAGTCTACTATGTCGAG 1686

Db 1301 CCTGTAAAGTACAGCGCCGGGAGACAAAGTGTGAGTGTAAAGTCACTATGTGGAG 1360
 QY 1687 ATGGGCTGAACCTGTAGCGGAGCAGCTGCCATTGACCGCTCTTACAGGACAAATGGGC 1746
 Db 1361 ATGGGCTGAACCTGTAGCGGAGCAGCTGCCATTGACCGCTCTTACAGGACAAATGGGC 1420
 QY 1747 AGTGCCATGACAGCGCAATGTGTGACCTCCACTTCACGATACACTGTGTGGGGTGT 1806
 Db 1421 AGTGCCATGACAGCGCAATGTGTGACCTCCACTTCACGATACACTGTGTGGGGTGT 1480
 QY 1807 TCCATCTAGCTCCCTGACCTGGGCGAGTAAAGCTGACCTTTGACAAAGCCAGAGGCGCT 1866
 Db 1481 TCCATCTAGCTCCCTGACCTGGGCGAGTAAAGCTGACCTTTGACAAAGCCAGAGGCGCT 1540
 QY 1867 GTGCCAACGAAGCTGGACATGGCACTTACCAACAGCTCTCTATGCCCCGAAG 1922
 Db 1541 GTGCCAACGAAGCTGGACATGGCACTTACCAACAGCTCTCTATGCCCCGAAG 1596

RESULT 14

AAC76373

ID AAC76373 standard; cDNA; 1377 BP.

XX AC AAC76373;
 XXXX 08-FEB-2001 (first entry)
 XXXX Human ORF1928 polynucleotide sequence SEQ ID NO:3855.
 XX

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antiparkinsonian; nootropic; neuroprotective; osteopathic;
 KW anticonvulsant; osteopathic; antidiabetic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antidiabetic;
 KW antiviral; antibacterial; cancer; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX 31-MAR-1999; 99US-0127607P.

XX 02-APR-1999; 99US-0127636P.

XX 05-APR-1999; 99US-0127728P.

XX 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB42164.

XX Novel nucleic acids and peptides derived from open reading frame X,

XX useful for treating e.g. cancers, proliferative disorders,

XX neurodegenerative disorders and cardiovascular disease.

XX Claim 5; Page 3006-3007; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,

XX which represent the human ORFX open reading frames 1 to 3161. The ORFX

CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiparkinsonian; nootropic; neuroprotective; osteopathic;
 CC anticonvulsant; antidiabetic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antirheumatic; antidiabetic;
 CC antiviral; antifungal; antineoplastic; antithyroid; and antianemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX

XX Sequence 1377 BP; 303 A; 398 C; 342 G; 330 T; 0 U; 4 Other;

Query Match 41.3%; Score 1347.4; DB 3; Length 1377;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1371; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1876 AAGCTGCGACCATGGCAACCTACAAACAGCTCTCTATGCCAGAGGCAAGTACCAACC 1935
 Db 2 AAGCTGCGACCATGGCAACCTACAAACAGCTCTCTATGCCAGAGGCAAGTACCAACC 61
 QY 1936 TGTGCTCAGCAGCTGGCTGGAGACCGGGCGGGTTGCCCTACCCACAGCTTCGCTCC 1995
 Db 62 TGTGCTCAGCAGCTGGCTGGAGACCGGGCGGGTTGCCCTACCCACAGCTTCGCTCC 121
 QY 1996 AGAAGCTGGCTCTGGTCTGGTGGATAGTGGACTATGGAGCTAGACCAACAGAGTG 2055
 Db 122 AGAAGCTGGCTCTGGTCTGGTGGATAGTGGACTATGGAGCTAGACCAACAGAGTG 181
 QY 2056 AAATGTGGAGTGTCTTCTGCTATCGGATGAAGATGTGAACCTGCACTGCAAGTGGGT 2115
 Db 182 AAATGTGGAGTGTCTTCTGCTATCGGATGAAGATGTGAACCTGCACTGCAAGTGGGT 241
 QY 2116 ATGTGGAGATGGCTTCTCATGCACTGGGAACCTGCTCAGTCTCTGATGTCTTCCCT 2175
 Db 242 ATGTGGAGATGGCTTCTCATGCACTGGGAACCTGCTCAGTCTCTGATGTCTTCCCT 301
 QY 2176 CACTCACAACTTCTGTGACGGAAGTGTGGGCTATTCCACAGCTCAGCTGAGGCGGTG 2235
 Db 302 CACTCACAACTTCTGTGACGGAAGTGTGGGCTATTCCACAGCTCAGCTGAGGCGGTG 361
 QY 2236 CATTTCTAGAACACTGACTGACCTGTCCATCCGCGGCACTCTTTGTGCCACAGACACA 2295
 Db 362 CATTTCTAGAACACTGACTGACCTGTCCATCCGCGGCACTCTTTGTGCCACAGACACA 421
 QY 2296 GTGGGCTGGGGAGAAATGAGACCTTGTGTGGCGGGGACATCGAGCACCACTCGCCAAATG 2355
 Db 422 GTGGGCTGGGGAGAAATGAGACCTTGTGTGGCGGGGACATCGAGCACCACTCGCCAAATG 481
 QY 2356 TCAGCATGTTTTTCTTACAATGACCTTGTCAATGACACC-ACCCTGCAACAGAGGTGGGA 2414
 Db 482 TCAGCATGTTTTTCTTACAATGACCTTGTCAATGACACCCTTCCAAACAGAGGTGGGA 541
 QY 2415 AGCAAGCTGTCTCATCTGCTGCGAGCAGGACCCACT--CCAAACGAGGAGACCAAGTTTG 2472
 Db 542 AGCAAGCTGTCTCATCTGCTGCGAGCAGGACCCACTNNCCAAACGAGGAGACCAAGTTTG 601
 QY 2473 TTGATGGAAGAGCAATTTCTGAGTGGGACATCTTTGCTTCAATGGGATCATTCATGTCA 2532
 Db 602 TTGATGGAAGAGCAATTTCTGAGTGGGACATCTTTGCTTCAATGGGATCATTCATGTCA 661
 QY 2533 TTTCCAGGCTTTAAAGACACCCCTGCCCCCGGTGACCTTGACCCACACTGGGTGGAG 2592
 Db 662 TTTCCAGGCTTTAAAGACACCCCTGCCCCCGGTGACCTTGACCCACACTGGGTGGAG 721

QY 2593 CAGGGATCTTCTTCCCATCATCTGGTGAAGTGGGGCTGTGGCTGGCTGCTTACTCT 2652
Db |||||
QY 722 CAGGGATCTTCTTCCCATCATCTGGTGAAGTGGGGCTGTGGCTGGCTGCTTACTCT 781
Db |||||
QY 2653 ACTTTTCGGATAAATCGGAGAAATCGGCTTCCAGCATTTTGGTTCGGAAGGACATTA 2712
Db |||||
QY 782 ACTTTTCGGATAAATCGGAGAAATCGGCTTCCAGCATTTTGGTTCGGAAGGACATTA 841
Db |||||
QY 2713 ATGTTGCGAGCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2772
Db |||||
QY 842 ATGTTGCGAGCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
Db |||||
QY 2773 CAACTCTGAGCTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2832
Db |||||
QY 902 CAACTCTGAGCTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 961
Db |||||
QY 2833 TTGAGGGCAATGACCCCTTTGAGGACACTGTGAGGGCTGGAGGGAGATGCCAGGCATCA 2892
Db |||||
QY 962 TTGAGGGCAATGACCCCTTTGAGGACACTGTGAGGGCTGGAGGGAGATGCCAGGCATCA 1021
Db |||||
QY 2893 CTCACCTGAGCTTGGGCGCATCACTGTGAATTTCTGAGCAGCTTCCCTTTTAGGAACGT 2952
Db |||||
QY 1022 CTCACCTGAGCTTGGGCGCATCACTGTGAATTTCTGAGCAGCTTCCCTTTTAGGAACGT 1081
Db |||||
QY 2953 AAAGTCTTTAAGCACTCAGAGCCATCACTCTCTGCTGATCTGGGGTTGTTT 3012
Db |||||
QY 1082 AAAGTCTTTAAGCACTCAGAGCCATCACTCTCTGCTGATCTGGGGTTGTTT 1141
Db |||||
QY 3013 CTGTGGGTGAGAGATGTTGTGTGTCGCCCAAGGAGTCTTCTCTGAGCCCTTTG 3072
Db |||||
QY 1142 CTGTGGGTGAGAGATGTTGTGTGTCGCCCAAGGAGTCTTCTCTGAGCCCTTTG 1201
Db |||||
QY 3073 GCTCTTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3132
Db |||||
QY 1202 GCTCTTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1261
Db |||||
QY 3133 AACTGTGATCTTCTTCCCTGTTAGATTGTAAGCCTCCCTTCTGATCCAGCCCTAG 3192
Db |||||
QY 1262 AACTGTGATCTTCTTCCCTGTTAGATTGTAAGCCTCCCTTCTGATCCAGCCCTAG 1321
Db |||||
QY 3193 CCAGTGTCTGACACAGGAACTGTGCAATAAAGTTTATGAAACAGAAAAA 3248
Db |||||
QY 1322 CCAGTGTCTGACACAGGAACTGTGCAATAAAGTTTATGAAACAGAAAAA 1377
Db |||||

RESULT 15
AAA57365
ID AAA57365 standard; DNA; 1259 BP.
XX
AC AAA57365;
XX
DT 03-OCT-2000 (first entry)
XX
DE DNA encoding a human hyaluronan-binding protein, designated BM-HABP.
XX
KW Hyaluronan-binding protein WF-HABP; OE-HABP; BM-HABP;
KW proliferative condition; metastasis; inflammation; ischemia;
KW host defence dysfunction; immune surveillance dysfunction; arthritis;
KW multiple sclerosis; autoimmunity; immune dysfunction; allergy; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FH 199..1257
CDS
FT /product= "hyaluronan-binding protein"
FT /*tag= a
FT /transl_except= (pos: 478..480, aa: Xaa)
FT /transl_except= (pos: 667..669, aa: Xaa)
FT /transl_except= (pos: 1105..1107, aa: Xaa)
FT /transl_except= (pos: 1138..1140, aa: Xaa)
FT /transl_except= (pos: 1156..1158, aa: Xaa)
FT /transl_except= (pos: 1168..1170, aa: Xaa)
FT /transl_except= (pos: 1171..1173, aa: Xaa)

/transl_except= (pos: 1249..1251, aa: Xaa)
/note= "Xaa is an unspecified amino acid; no termination
codon given"

WO200039166-A1.
06-JUL-2000.
20-DEC-1999; 99WO-US030462.
23-DEC-1999; 98US-0113871P.
(HUMA-) HUMAN GENOME SCI INC.
(AMNA-) AMERICAN NAT RED CROSS.
Hastings GA, Liau G, Tsifrina E;
WPI; 2000-452376/39.
P-PSDB; AAY93913.
New hyaluron-binding proteins, known as full-length WF-HABP, WF-HABP, OE-
HABP and BM-HABP, useful for treating proliferative conditions,
metastasis, inflammation, ischemia, arthritis and multiple sclerosis.
Claim 2; Fig 4A-B; 457pp; English.
The present sequence encodes a hyaluronan-binding protein. The
specification describes four hyaluronan-binding proteins, known as WF-
HABP, WF-HABP, OE-HABP, and BM-HABP. The polypeptides are useful for
treating diseases such as proliferative conditions, metastasis,
inflammation, ischemia, host defence dysfunction, immune surveillance
dysfunction, arthritis, multiple sclerosis, autoimmunity, immune
dysfunction and allergy

Sequence 1259 BP; 295 A; 348 C; 346 G; 258 T; 0 U; 12 Other;
Query Match 37.7%; Score 1227.4; DB 3; Length 1259;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1234; Conservative 2; Mismatches 23; Indels 0; Gaps 0;

QY 1437 GGNATCACATGCACAGTGTGGATTCTGCAACAGACAGCAACGGGGCTGTGCAAGGTG 1496
Db 1 GGNATCACATGCACAGTGTGGATTCTGCAACAGACAGCAACGGGGCTGTGCAAGGTG 60
QY 1497 GCCAGATGTCCTCCAGAAAGGCGACGAAGTCTCTCGAGCTGCCAGAGGGATCAAGGG 1556
Db 61 GCCAGATGTCCTCCAGAAAGGCGACGAAGTCTCTCGAGCTGCCAGAGGGATCAAGGG 120
QY 1557 GACGGGACAGCTGCAAGAGATAGACCTCTGTGAGAGCGCTTACCGAGGCTGTAC 1616
Db 121 GACGGGACAGCTGCAAGAGATAGACCTCTGTGAGAGCGCTTACCGAGGCTGTAC 180
QY 1617 GAGCAGCCACCTGTAAGATGACAGCGCCGGGCAAGCACAAGTGTGAGTGTAAAGTAC 1676
Db 181 GAGCAGCCACCTGTAAGATGACAGCGCCGGGCAAGCACAAGTGTGAGTGTAAAGTAC 240
QY 1677 TATGTCCGAGATGGCTGAACTGTGAGCCGAGAGCTGCCATTTGACCCGCTTACAG 1736
Db 241 TATGTCCGAGATGGCTGAACTGTGAGCCGAGAGCTGCCATTTGACCCGCTTACAG 300
QY 1737 GACAAATGGGAGTGCCTGAGAGCGCAAGTGTGAGCTTCCAGGATACCACT 1796
Db 301 GACAAATGGGAGTGCCTGAGAGCGCAAGTGTGAGCTTCCAGGATACCACT 360
QY 1797 GTTGGGCTGTTCATCTACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1856
Db 361 GTTGGGCTGTTCATCTACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 420
QY 1857 AGAGAGCCCTGTGCAACAGAGCTGCAAGTGTGAGCTTCCAGGATACCACTCCCTATGCC 1916
Db 421 AGAGAGCCCTGTGCAACAGAGCTGCAAGTGTGAGCTTCCAGGATACCACTCCCTATGCC 480
QY 1917 CAGNAGGCCAAGTACCACTGCTGCTCAGAGGCTGGTGGAGAGACCGGCGGGTTCCTTAC 1976

[illegible]

Search completed: June 21, 2004, 10:47:18
Job time : 1193 secs

